



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 103887

TO: Vanessa L Ford

Location:

Art Unit: 1645

Tuesday, September 16, 2003

Case Serial Number: 09596101

From: Edward Hart

Location: Biotech-Chem Library

CM1-6B02

Phone: 305-9203

edward.hart@uspto.gov

Search Notes

Examiner Ford,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

From: Chan, Christina
Sent: Monday, September 15, 2003 8:51 AM
T : Ford, Vanessa; STIC-Biotech/ChemLib
Subject: RE: In re:09596101 Sequence search

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
308-3973
CM-1, 9B19

-----Original Message-----

From: Ford, Vanessa
Sent: Saturday, September 13, 2003 5:15 PM
To: Chan, Christina
Subject: In re:09596101 Sequence search

Please search SEQ ID NO:1 and 3.

Please include interference searches. Please rush.

Vanessa L. Ford
Biotechnology Patent Examiner
Office: CM1 8A16
Mailbox: CM1 8E12
Phone: 703.308.4735
Art Unit:1645

Edward Hari
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 9/16/03
Date Completed: _____
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: ✓
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: 102
WWW/Internet: _____
Other (specify): _____

Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. The Pending databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches. Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions .rnpm and .rnpn. Searches run against the Amino Acid Pending database produce two sets of results, with the extensions .rapm and .rapn.

Because they contain data that is confidential, the results of Pending database searches should not be left in the case.



STIC SEARCH RESULTS

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or contact:

Mary Hale, Information Branch Supervisor
308-4258, CM1-1E01

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art found, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art not found:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library CM1 - Circ. Desk



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:14:52 ; Search time 3.96222 Seconds

(without alignments)
520.781 Million cell updates/sec

Title: US-09-596-101c-1

Perfect score: 64
Sequence: 1 SEEDIDITETGR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	13	20	AAV24915
2	64	100.0	384	20	AAV24914
3	50	78.1	306	17	AAV24914
4	45	70.3	294	22	AAV24914
5	44	68.8	263	18	AAV24914
6	44	68.8	303	18	AAV24914
7	44	68.8	435	18	AAV24914
8	43	67.2	384	22	AAV24914
9	40	62.5	185	22	AAV24914

10	40	62.5	205	22	AAV45016
11	40	62.5	254	22	ABG04191
12	40	62.5	501	22	ABG04189
13	40	62.5	575	23	ABP27730
14	40	62.5	575	23	ABP29826
15	40	62.5	1080	22	ABG04194
16	40	62.5	1342	22	AAU34845
17	40	62.5	1343	22	AAU35444
18	40	62.5	1380	22	AAU38450
19	40	62.5	1394	24	ABP80700
20	39	60.9	261	12	AAV11599
21	39	60.9	297	17	AAV88406
22	39	60.9	540	22	ABP3776
23	39	60.9	665	23	ABV3493
24	38	59.4	223	21	AAV32318
25	38	59.4	1195	22	ABG24229
26	37	57.8	112	22	ABG17370
27	37	57.8	275	21	AAV44125
28	37	57.8	324	21	AAV44124
29	37	57.8	336	21	AAV44123
30	37	57.8	1292	22	ABV70537
31	36	56.2	222	22	AAE00115
32	36	56.2	352	22	ABV58666
33	36	56.2	419	22	ABV70788
34	36	56.2	441	22	AAV31857
35	36	56.2	578	23	ABV30909
36	36	56.2	606	23	ABV47745
37	36	56.2	659	22	AAU34917
38	36	56.2	769	22	ABG13410
39	36	56.2	1149	22	AAV97640
40	36	56.2	1188	23	ABP25602
41	36	56.2	1191	23	ABP25601
42	36	56.2	1194	20	AAV91071
43	36	56.2	1194	21	AAV49432
44	36	56.2	1194	22	AAV97649
45	36	56.2	1194	23	ABV56662

ALIGNMENTS

RESULT 1	AAV24915
ID	AAV24915 standard; peptide: 13 AA.
XX	XX
AC	AAV24915;
XX	XX
DT	25-AUG-1999 (first entry)
XX	XX
DE	Eisenia foetida coelomic cytolytic factor 1 peptide.
XX	XX
KW	Eisenia foetida; coelomic cytolytic factor 1; CCF-1; cancer;
KW	typhlococcal infection; bacterial infection; tumour therapy;
KW	inflammation; immunology.
XX	XX
OS	Eisenia foetida.
XX	XX
PN	WC09931229-A2.
XX	XX
PD	24-JUN-1999.
XX	XX
PF	16-DEC-1998; 98WO-EP08169.
XX	XX
PR	17-DEC-1997; 97EP-0203974.
XX	XX
PA	(VLAAR-) VLAAR INTERUNIVERSITAIR INST BIOTECHNOG.
XX	XX
PI	De Baetselier P;
XX	XX
DR	WPI; 1999-385905/32.
XX	XX
PT	Eisenia foetida polypeptides derived from coelomic cytolytic factor 1

Proionibacterium
Novel human diago
Novel human diago
Streptococcus poly
Streptococcus poly
Novel human diago
E. coli cellular p
Haemophilus influ
Salmonella typhi c
N. gonorrhoeae ami
Beta-1,3-glucanase
Trichoderma harzia
Candida albicans e
Drosophila melano
Corn beta-carotene
Novel human diago
Novel human diago
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Drosophila melano
Bacillus lichenifo
Drosophila melano
N. magdali bacter
C. glutamicum prote
Mycobacterium tube
Listeria monocyco
Enterococcus faeca
Novel human diago
Araf-1X(L)Delta2-10
Streptococcus poly
Streptococcus poly
Apoptosis inducer
Human full-length
Wild type Araf-1 p
Human apoptotic pr

PS Claim 1; Page 45; 49pp; English.

XX The present sequence represents a *Eisenia foetida* coelomic cytolytic

CC factor 1 (CCF-1) peptide. The CCF-1 protein has antiparasitic,

CC antibacterial and anti-inflammatory activity. Recombinant coelomic

CC cytolytic factor 1 (rCCF-1) is trypanolytic for the African trypanosome

CC *Trypanosoma brucei* in a dose-dependent manner. The trypanolytic activity

CC of rCCF-1 can be inhibited by anti-CCF-1 and anti-tumour necrosis factor

CC (TNF)/TNF monoclonal antibodies. Furthermore, N,N'-diacetylchitobiose

CC inhibits potentially trypanolytic activity of rCCF-1. These data

CC corroborate the findings that CCF-1 shares a trypanolytic, lectin-like

CC domain with TNF-alpha. CCF-1 is useful to treat trypanosomal or

CC bacterial infections or cancer. The proteins and peptides are also

CC useful in tumour therapy, inflammation and other areas of immunology.

CC The annelid peptide is derived from a 42 kDa cytolytic protein named

CC CCF-1 that binds lipopolysaccharide and beta-1,3-glucan. The factor

CC resembles the vertebrate tumour necrosis factor-alpha (TNF-alpha), and

CC may be used as an alternative for TNF-alpha.

XX Sequence 13 AA;

SO

Query Match 100.0%; Score 64; DB 20; Length 13;

Best Local Similarity 100.0%; Pred. No. 1.8e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEIDIIETIGNR 13

DB 1 SGEIDIIETIGNR 13

RESULT 2

AAY24914

ID AAY24914 standard; Protein; 384 AA.

AC AAY24914;

XX

XX 25-AUG-1999 (first entry)

DT

XX

DE *Eisenia foetida* coelomic cytolytic factor 1 protein.

XX

XX *Eisenia foetida*; coelomic cytolytic factor 1; CCF-1; cancer;

KW trypanosomal infection; bacterial infection; tumour therapy;

KM inflammation; immunology.

XX

OS *Eisenia foetida*.

XX

XX Key Location/Qualifiers

FT Peptide 1..17

FT /label= signal

FT 18..384

FT Protein /label= CCF-1

XX

PN W09931229-A2.

XX

XX 24-JUN-1999.

PD

XX

XX 16-DEC-1998; 98WO-EP08169.

PF

XX

XX 17-DEC-1997; 97EP-0203974.

PR

XX

XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

PA

XX

PI De Baetselier P;

PI

XX

DR WPI: 1999-385905/32.

DR N-PSDB; AAX83611.

XX

XX *Eisenia foetida* polypeptides derived from coelomic cytolytic factor 1

PT

XX

PS Claim 2; Page 48-49; 49pp; English.

XX

CC The present sequence represents the *Eisenia foetida* coelomic cytolytic

CC factor 1 (CCF-1). The protein has antiparasitic, antibacterial and

CC anti-inflammatory activity. Recombinant coelomic cytolytic factor 1

CC (rCCF-1) is trypanolytic for the African trypanosome *Trypanosoma brucei*

CC in a dose-dependent manner. The trypanolytic activity of rCCF-1 can be

CC inhibited by anti-CCF-1 and anti-tumour necrosis factor (TNF)/TNF

CC monoclonal antibodies. Furthermore, N,N'-diacetylchitobiose inhibits

CC potentially trypanolytic activity of rCCF-1. These data corroborate the

CC findings that CCF-1 shares a trypanolytic, lectin-like domain with

CC TNF-alpha. CCF-1 is useful to treat trypanosomal or bacterial infections

CC or cancer. The proteins and peptides are also useful in tumour therapy,

CC inflammation and other areas of immunology. The annelid peptide is

CC derived from a 42 kDa cytolytic protein named CCF-1 that binds

CC lipopolysaccharide and beta-1,3-glucan. The factor resembles the

CC vertebrate tumour necrosis factor-alpha (TNF-alpha), and may be used as

CC an alternative for TNF-alpha.

XX Sequence 384 AA;

SO

Query Match 100.0%; Score 64; DB 20; Length 384;

Best Local Similarity 100.0%; Pred. No. 0.00092;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEIDIIETIGNR 13

DB 178 SGEIDIIETIGNR 190

RESULT 3

AAR97362

ID AAR97362 standard; Protein; 306 AA.

AC AAR97362;

XX

XX 03-OCT-1996 (first entry)

DT

XX

DE *Oerskovia beta*-1,3-glucanase.

XX

XX *Beta*-1,3-glucanase; *Cellulomonas cellulans*; *Bacillus subtilis*;

KW lytic enzyme; beta-glucan degradation; cell wall lysis;

KM pigment; colorant; flavour; yeast extract; protoplast.

XX

OS *Oerskovia xanthineolytica* strain LUG109 (DSM 10297).

XX

XX Key Location/Qualifiers

FT Peptide 1..35

FT /label= sig_peptide

FT 36..63

FT /label= Pro_peptide

FT Protein /label= Mat_protein

XX

PN W09612013-A1.

XX

XX 25-APR-1996.

PD

XX

XX 16-OCT-1995; 95WO-DK00414.

PF

XX

XX 14-OCT-1994; 94DK-0001192.

PR

XX

XX (NOVO) NOVO-NORDISK AS.

PA

XX

PI Asenjo JA, Diers I, Ferrer P, Halkier T, Hedegaard L;

PI Savva D;

PI

XX

DR WPI: 1996-222000/22.

DR N-PSDB; AAT29043.

XX

XX DNA construct encoding enzyme with beta-1,3-glucanase activity -

PT useful for modifying or degrading beta-glucan contg. material and in

PT the prepn. of e.g. food colourants, flavourings and yeast extracts

XX

PS Claim 1; Page 42-43; 60pp; English.

XX

CC A novel beta-1,3-glucanase (AAR97362) from *Oerskovia xanthineolytica*

CC LLG109 is useful for degrading or modifying beta-glucan contg.
 CC material. Its amino acid sequence was deduced from a genomic
 CC DNA sequence (AAF29043) isolated from an O. xanthineolytica library.
 CC Recombinant beta-1,3-glucanase can be produced on a large scale
 CC using transformed host cells, esp. Bacillus subtilis DN1885 or
 CC TOC46. Protease-free beta-1,3-glucanase can be obtd. that is useful
 CC for lysing fungal cell walls, allowing recovery of intracellular
 CC proteins. The enzyme is also useful for the prepn. of protoplasts
 CC and for the prodn. of pigments, colorants, flavours, yeast extract
 CC and pharmaceuticals.
 XX
 SQ Sequence 306 AA;
 Query Match 78.1%; Score 50; DB 17; Length 306;
 Best Local Similarity 75.0%; Pred. No. 0.33;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SGEIDITETIGN 12
 DB 180 SGEIDITETIGN 191
 RESULT 4
 ID AAB52463 standard; protein; 294 AA.
 XX AAB52463;
 NC AAB52463;
 XX
 DT 23-FEB-2001 (first entry)
 XX
 DE Mycobacterium tuberculosis secreted protein #28.
 XX
 KM Mycobacterium tuberculosis secreted protein; MTSP; vaccine.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN W020006143-A1.
 XX
 PD 09-NOV-2000.
 XX
 PF 04-MAY-2000; 2000MO-US12197.
 XX
 PR 04-MAY-1999; 99US-0132479.
 PR 04-MAY-1999; 99US-0132503.
 XX
 PA (PUBL-) PUBLIC HEALTH RES INST NEW YORK.
 XX
 PI Genaro ML, Gomez MJ;
 XX
 DR WPI; 2001-007151/01.
 XX
 PT Novel Mycobacterium tuberculosis secreted polypeptides and
 PT polynucleotides useful in diagnosis, treatment and prophylaxis of
 PT tuberculosis -
 XX
 PS Claim 11; Fig 1; 60pp; English.
 XX
 CC The present invention relates to Mycobacterium tuberculosis secreted
 CC proteins (MTSP), where the polypeptide has M. tuberculosis specific
 CC antigenic and immunogenic properties. Compositions of the invention may
 CC be useful for diagnosing Mycobacterium tuberculosis infection and as a
 CC vaccine against M. tuberculosis infection.
 CC
 SQ Sequence 294 AA;
 Query Match 70.3%; Score 45; DB 22; Length 294;
 Best Local Similarity 75.0%; Pred. No. 2.9;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SGEIDITETIGN 12
 DB 169 SGEIDITETIGN 180

RESULT 5
 AAW29455
 ID AAW29455 standard; protein; 263 AA.
 XX
 AC AAW29455;
 XX
 DT 14-APR-1998 (first entry)
 XX
 DE Oerskovia xanthineolytica mature beta-1,3-glucanase.
 XX
 KM Beta-1,3-glucanase; lytic enzyme; yeast; beta glucan degradation;
 KM fungal cell wall; intracellular product; purification; protoplast.
 OS Oerskovia xanthineolytica LLG109 (DSM 10297).
 XX
 FH Key Location/Qualifiers
 FT CDS 23..955 /*tag= a
 FT sig_peptide 23..120 /*tag= b
 FT mat_peptide 164..952 /*tag= c
 FT
 XX
 PN W09739114-A1.
 XX
 PD 23-OCT-1997.
 XX
 PE 14-APR-1997; 97WO-DK00160.
 XX
 PR 23-AUG-1996; 96DK-0000885.
 PR 12-APR-1996; 96DK-0000427.
 XX
 PA (NOVO) NOVO-NORDISK AS.
 XX
 PI Diers I, Ferrer P, Halkier T, Hedegaard L;
 XX
 DR WPI; 1997-526451/48.
 DR N-PSDB; AAT89155.
 XX
 PT New isolated beta-1,3-glucanase enzyme - obtained from Oerskovia
 PT xanthineolytica, used particularly for the lysis of microbial cells
 PT for obtaining desirable products
 XX
 PS Example 2; Page 35-36; 64pp; English.
 XX
 CC This polypeptide comprises a novel Oerskovia xanthineolytica (OX)
 CC enzyme that exhibits beta-1,3-glucanase (BG) activity. Its amino
 CC acid sequence was deduced from an isolated genomic DNA sequence
 CC (see AAT89155). Claimed DNA constructs that encode the novel BG (see
 CC also AAW29456 for corrected sequence), a mannose binding domain (see
 CC AAW29458) or a full-length enzyme, i.e. BG with mannose binding
 CC domain (see AAW29456), can be used to produce recombinant BG
 CC polypeptides, with or without a mannose binding domain, in fungal
 CC or bacterial host cells. BG polypeptides are used for the
 CC degradation or modification of beta-glucan containing material,
 CC especially for the gentle lysis of microbial cell walls, thereby
 CC enabling recovery of desirable intracellular products with a
 CC reduced amount of contaminants. They can also be used for the
 CC production of e.g. pigments, colorants, flavours, yeast
 CC extracts, pharmaceuticals, food or feed compositions, and to
 CC prepare protoplasts for use in fusion, transformation and cloning
 CC studies.
 XX
 SQ Sequence 263 AA;
 Query Match 68.8%; Score 44; DB 18; Length 263;
 Best Local Similarity 72.7%; Pred. No. 3.9;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SGEIDITETIG 11
 DB 116 SGEIDITETIG 126

RESULT 6

AAW29457

ID AAW29457 standard; Protein: 303 AA.

AC AAW29457;

DT 14-APR-1998 (first entry)

DE Oerskovia xanthineolytica beta-1,3-glucanase.

KM Beta-1,3-glucanase; lytic enzyme; yeast; beta glucan degradation;
KW fungal cell wall; intracellular product; purification; protoplast.

OS Oerskovia xanthineolytica LIG109 (DSM 10297).

FH Key Location/Qualifiers

FT Peptide 1..52

FT Protein /label= Sig_peptide

FT Protein 53..303

FT Protein /label= Mat_protein

PN WO9739114-A1.

PD 23-OCT-1997.

PF 14-APR-1997; 97WO-DK00160.

PR 23-AUG-1996; 96DK-0000885.

PR 12-APR-1996; 96DK-0000427.

XX (NOVO) NOVO-NORDISK AS.

PI Diers I, Ferrer P, Halkier T, Hedegaard L;

DR WPI: 1997-526451/48.

DR N-PSDB; AAT89157.

PT New isolated beta-1,3-glucanase enzyme - obtained from Oerskovia
PT xanthineolytica, used particularly for the lysis of microbial cells
PT for obtaining desirable products

PS Example 2; Page 42-43; 64pp; English.

CC This sequence comprises the polypeptide precursor of a novel
CC Oerskovia xanthineolytica enzyme that exhibits beta-1,3-glucanase
CC (BG) activity. Its amino acid sequence was deduced from an
CC isolated genomic DNA sequence (see AAT89157). Claimed DNA constructs
CC that encode the novel BG (see also AAW29455), a mannose binding
CC domain (see AAW29458) or a full-length enzyme, i.e. BG with mannose
CC binding domain (see AAW29456), can be used to produce recombinant BG
CC polypeptides, with or without a mannose binding domain, in fungal
CC or bacterial host cells. BG polypeptides are used for the
CC degradation or modification of beta-glucan containing material,
CC especially for the gentle lysis of microbial cell walls, thereby
CC enabling recovery of desirable intracellular products with a
CC reduced amount of contaminants. They can also be used for the
CC production of e.g. pigments, colourants, flavourants, yeast
CC extracts, pharmaceuticals, food or feed compositions, and to
CC prepare protoplasts for use in fusion, transformation and cloning
CC studies.

SQ Sequence 303 AA;

Query Match 68.8%; Score 44; DB 18; Length 303;

Best Local Similarity 72.7%; Pred. No. 4.6;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGEIDIIETIG 11

DB 169 SGEIDIMENVG 179

RESULT 7

AAW29456

ID AAW29456 standard; Protein: 435 AA.

AC AAW29456;

DT 14-APR-1998 (first entry)

DE Oerskovia xanthineolytica beta-1,3-glucanase.

KM Beta-1,3-glucanase; lytic enzyme; yeast; beta glucan degradation;
KW fungal cell wall; intracellular product; purification; protoplast.

OS Oerskovia xanthineolytica LIG109 (DSM 10297).

FH Key Location/Qualifiers

FT Peptide 1..52

FT Protein /label= Sig_peptide

FT Protein 53..435

FT Protein /label= Mat_protein

FT Protein 304..435

FT Protein /label= Mannose-binding_domain

PN WO9739114-A1.

PD 23-OCT-1997.

PF 14-APR-1997; 97WO-DK00160.

PR 23-AUG-1996; 96DK-0000885.

PR 12-APR-1996; 96DK-0000427.

XX (NOVO) NOVO-NORDISK AS.

PI Diers I, Ferrer P, Halkier T, Hedegaard L;

DR WPI: 1997-526451/48.

DR N-PSDB; AAT89156.

PT New isolated beta-1,3-glucanase enzyme - obtained from Oerskovia
PT xanthineolytica, used particularly for the lysis of microbial cells
PT for obtaining desirable products

PS Example 2; Page 39-40; 64pp; English.

CC This sequence comprises the polypeptide precursor of a novel
CC Oerskovia xanthineolytica enzyme that exhibits beta-1,3-glucanase
CC (BG) activity and which includes a mannose binding domain (MBD).
CC Its amino acid sequence was deduced from an isolated genomic DNA
CC sequence (see AAT89156). Claimed DNA constructs that encode the
CC novel BG lacking a MBD (see AAW29455 and AAW29457), a MBD (see
CC AAW29458), or the full-length enzyme can be used to produce recombinant
CC BG polypeptides, with or without a mannose binding domain, in fungal
CC or bacterial host cells. BG polypeptides are used for the
CC degradation or modification of beta-glucan containing material,
CC especially for the gentle lysis of microbial cell walls, thereby
CC enabling recovery of desirable intracellular products with a
CC reduced amount of contaminants. They can also be used for the
CC production of e.g. pigments, colourants, flavourants, yeast
CC extracts, pharmaceuticals, food or feed compositions, and to
CC prepare protoplasts for use in fusion, transformation and cloning
CC studies.

SQ Sequence 435 AA;

Query Match 68.8%; Score 44; DB 18; Length 435;

Best Local Similarity 72.7%; Pred. No. 7;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGEIDIIETIG 11

DB 169 SGEIDIMENVG 179

RESULT 8
 AAB99272
 ID AAB99272 standard; protein; 384 AA.
 AC AAB99272;
 XX
 DT 11-SEP-2001 (first entry)
 DE Bacillus circulans beta-1,3-glucanase.
 XX
 KM beta-1,3-glucanase; enzyme; stockbreeding; bread manufacture; brewing.
 XX
 OS Bacillus circulans.
 XX
 PN JP2001120280-A.
 XX
 PD 08-MAY-2001.
 XX
 PF 01-NOV-1999; 99JP-0311073.
 XX
 PR 01-NOV-1999; 99JP-0311073.
 XX
 PA (MEIJI) MEIJI SEIKA KASHA LTD.
 XX
 DR WPI: 2001-364762/38.
 XX
 DR N-PSDB; AAH46359, AAH46360.
 XX
 PT Beta-1,3-glucanase and beta-1,3-glucan-binding peptide and their genes
 PS Claim 1; Page 8-9; 12pp; Japanese.
 XX
 CC The present sequence is beta-1,3-glucanase from Bacillus circulans. The
 CC protein can be used in various industrial fields such as stockbreeding,
 CC cake and bread manufacture and brewing.
 XX
 SQ Sequence 384 AA;
 Query Match 67.2%; Score 43; DB 22; Length 384;
 Best Local Similarity 66.7%; Pred. No. 9.4;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SEIDITION 12
 DB 118 SGEIDIMERVN 129

RESULT 9
 ABG04199
 ID ABG04199 standard; Protein; 185 AA.
 AC ABG04199;
 XX
 DT 13-FEB-2002 (first entry)
 DE Novel human diagnostic protein #4190.
 XX
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 XX
 PR 23-AUG-2000; 2000US-0649167.
 XX

PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 XX
 DR N-PSDB; AAS68386.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID No 34558; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at http://wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 185 AA;
 Query Match 62.5%; Score 40; DB 22; Length 185;
 Best Local Similarity 58.3%; Pred. No. 15;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 2 GEIDITIONR 13
 DB 103 GEVDIDIDHLCNR 114

RESULT 10
 AAU45016
 ID AAU45016 standard; Protein; 205 AA.
 AC AAU45016;
 XX
 DT 27-FEB-2002 (first entry)
 DE Propionibacterium acnes immunogenic protein #5912.
 XX
 KM SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KM urethritis; endophthalmitis; bone, joint, central nervous system; ELISA;
 KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KM dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 XX
 PR 02-JUN-2000; 2000US-208841P.
 XX
 PR 07-JUL-2000; 2000US-216747P.
 XX

XX (CORI-) CORIXA CORP.
 PA
 XX
 PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI: 2001-616774/71.
 DR N-PSDB: AAS59524.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID No 6211; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), warts and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 205 AA:
 XX
 OY 2 GEIDIIETIGNR 13
 1 |||:|:|:|
 Db 66 GLIDIVESYCTR 77
 XX
 RESULT 11
 ABG04191
 ID ABG04191 standard; Protein; 254 AA.
 XX
 AC ABG04191:
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #4182.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEO INC.
 XX

PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB: AAS68378.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID No 34550; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 254 AA:
 XX
 OY 2 GEIDIIETIGNR 13
 11:1:|:|:|
 Db 180 GEVDEIDHLGNR 191
 XX
 RESULT 12
 ABG04189
 ID ABG04189 standard; Protein; 501 AA.
 XX
 AC ABG04189:
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #4180.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEO INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX

PRI: 2001-639362/73.
N-PSTB: AAS68376.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -

Claim 20; SEQ ID NO 34548; 103bp; English.

The invention relates to isolated polynucleotide (I) and poly peptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG3037 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 501 AA;

Query	March	62.5%	Score 40;	DB 22;	Length 501;
Best Local Similarity	58.38;	Pred. No. 48;			
Matches	7;	Conservative	3;	Mismatches	
			2;	Indels	0; Gaps 0;
OY	2 GEIDIIEITGNR 13	: : :			
Db	223 GEVDIDHLCNR 234				
RESULT 13					
ABP27730					
ID	ABP27730 standard; Protein: 575 AA.				
XX					
AC	ABP27730;				
XX					
DT	02-JUL-2002 (first entry)				
XX					
DE	Streptococcus polypeptide SEQ ID NO 4636.				
XX					
KW	Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;				
KW	group A streptococcus; Streptococcus pyogenes; antibacterial;				
KW	antiinflammatory; infection; vaccine; meningitis; gene therapy.				
XX					
OS	Streptococcus agalactiae.				
XX					
PN	WO200234771-A2.				
XX					
PD	02-MAY-2002.				
XX					
PF	29-OCT-2001; 2001WO-GB04789.				
XX					
PR	27-OCT-2000; 2000GB-0026333.				
PR	24-NOV-2000; 2000GB-0028727.				
PR	07-MAR-2001; 2001GB-0005640.				
XX					
PA	(CHIR-) CHIRON SPA.				
PA	(GENO-) INST GENOMIC RES.				
XX					
TI	Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;				

PI Tettelin H;
XX
XX WPI; 2002-352536/38.
DR N-PSDB; ABN68361.
XX
XX
PI New Streptococcus protein for the treatment or prevention of infection
PI or disease caused by Streptococcus bacteria, such as meningitis, and
PI for detecting a compound that binds to the protein -

Claim 1; Page 3626; 4525pp; English.

The invention relates to a protein (ABP25413-ABP30895) from group B streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS (Streptococcus pyogenes), comprising one of 5483 sequences (51), given in the specification. The proteins have antibacterial and anti-inflammatory activity. (1), nucleic acids encoding (1), ABN6604-ABN71526 and antibodies that bind (1) are used in the manufacture of medicaments for the treatment or prevention of infection or disease caused by Streptococcus bacteria, particularly *S. agalactiae* and *S. pyogenes*. Nucleic acids encoding (1) are used to detect Streptococcus in a biological sample. (1) is used to determine whether a compound binds to (1). A composition comprising (1) or a nucleic acid encoding (1), may be used as a vaccine or diagnostic composition. The disease caused by Streptococcus that is prevented or treated may be meningitis. Nucleic acid encoding (1) may be used to recombinantly produce (1) and may be used in gene therapy. Antibodies to (1) are used for affinity chromatography, immunoassays, and distinguishing/identifying Streptococcus proteins.

SQ Sequence 575 AA;

Query Match	62.5%	Score 40	DB 23	Length 575
Best Local Similarity	66.7%	Pred. No. 56		
Matches	8	Conservative	1	Mismatches 3
				Indels 0
				Gaps 0

```
QY      1 SGEDIETIGN 12
Db      27 SGLFDFIQTIGN 38
```

RESULT 14	
ABP29826	standard; Protein; 575 AA.
ID	ABP29826
XX	
AC	ABP29826;
XX	
DT	02-JUL-2002 (first entry)
XX	
DE	Streptococcus polypeptide SEQ ID NO 8828.
XX	
KM	Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
XX	
KM	group A streptococcus; Streptococcus pyogenes; antibacterial;
XX	
KM	antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX	
OS	Streptococcus agalactiae.
XX	
PN	WO200234771-A2.
XX	
PD	02-MAY-2002.
XX	
PF	29-OCT-2001; 2001WO-GB04789.
XX	
PR	27-OCT-2000; 2000GB-0026333.
XX	
PR	24-NOV-2000; 2000GB-0028727.
XX	
PR	07-MAR-2001; 2001GB-0005640.
XX	
PA	(CHIR-) CHIRON SPA.
XX	
PA	(GENO-) INST GENOMIC RES.
XX	
PI	Telford J, Maignant V, Margalit Ros YI, Grandi G, Fraser C;
XX	
PI	Tetelin H;
XX	
WPI:	2002-352536/38.

DR N-PSDB; ABN70457.
XX
PT New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
PS Claim 1; Page 4000; 4525pp; English.
XX
CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (II), nucleic acids encoding (I), ABN60044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
SQ Sequence 575 AA;
Query Match 62.5%; Score 40; DB 23; Length 575;
Best Local Similarity 66.7%; Pred. No. 56;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 SEIDIIETIGN 12
DB 27 SGLFDFIQTIGN 38
II I I I I I I
I I I I I I
RESULT 15
ABG04194
ID ABG04194 standard; Protein: 1080 AA.
XX
AC ABG04194;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #4185.
XX
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PE 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Dmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS68381.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX

PS Claim 20; SEQ ID No 34553; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1080 AA;
Query Match 62.5%; Score 40; DB 22; Length 1080;
Best Local Similarity 58.3%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 2 GEIDIIETIGNR 13
DB 219 GEVDIDIDHLGNR 230
II I I I I I I
I I I I I I
Search completed: September 16, 2003, 11:25:53
Job time : 4.96222 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:23:52 ; Search time 1.37531 Seconds
(without alignments)
399.938 Million cell updates/sec

Title: US-09-596-101C-1
Perfect score: 64
Sequence: 1 SGEIDIEITGNR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCtUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	78.1	306	2	US-08-824-707-2 Sequence 2, Appl 1
2	44	68.8	263	3	US-09-159-106-2 Sequence 2, Appl 1
3	44	68.8	303	3	US-09-159-106-13 Sequence 13, Appl 1
4	44	68.8	435	3	US-09-159-106-11 Sequence 11, Appl 1
5	40	62.5	816	4	US-09-266-225D-12 Sequence 12, Appl 1
6	40	62.5	1380	4	US-09-328-352-8132 Sequence 8132, Ap
7	39	60.9	226	2	US-08-712-072C-4 Sequence 4, Appl 1
8	39	60.9	232	2	US-08-712-072C-4 Sequence 4, Appl 1
9	39	60.9	232	2	US-08-712-072C-4 Sequence 4, Appl 1
10	39	60.9	232	2	US-08-712-072C-4 Sequence 4, Appl 1
11	37	57.8	262	1	US-08-392-828C-37 Sequence 37, Appl 1
12	37	57.8	262	1	US-08-392-828C-37 Sequence 37, Appl 1
13	37	57.8	262	1	US-08-392-828C-37 Sequence 37, Appl 1
14	36	56.2	321	2	US-08-712-072C-3 Sequence 3, Appl 1
15	36	56.2	1194	3	US-09-092-508-2 Sequence 2, Appl 1
16	36	56.2	1194	3	US-09-092-508-2 Sequence 2, Appl 1
17	36	56.2	1194	3	US-09-092-508-2 Sequence 2, Appl 1
18	36	56.2	1194	3	US-09-092-508-2 Sequence 2, Appl 1
19	36	56.2	1205	3	US-09-092-508-16 Sequence 16, Appl 1
20	35	54.7	233	4	US-09-435-115-16 Sequence 16, Appl 1
21	35	54.7	233	4	US-09-435-115-16 Sequence 16, Appl 1
22	35	54.7	239	4	US-09-328-352-6383 Sequence 25247, A
23	35	54.7	430	1	US-08-476-008-44 Sequence 29159, A
24	35	54.7	430	1	US-08-476-008-44 Sequence 29159, A
25	35	54.7	430	1	US-08-476-008-44 Sequence 29159, A
26	35	54.7	430	1	US-08-476-008-44 Sequence 29159, A
27	35	54.7	430	1	US-08-476-008-44 Sequence 29159, A

28	35	54.7	746	4	US-09-252-991A-18587 Sequence 18587, A
29	35	54.7	1151	4	US-09-252-991A-21328 Sequence 21328, A
30	34.5	53.9	448	4	US-09-252-991A-21043 Sequence 21043, A
31	34	53.1	141	4	US-09-328-352-6280 Sequence 6280, Ap
32	34	53.1	195	4	US-09-252-991A-27440 Sequence 27440, A
33	34	53.1	342	4	US-09-632-947B-2 Sequence 2, Appl 1
34	34	53.1	482	3	US-09-457-046B-63 Sequence 63, Appl 1
35	34	53.1	582	4	US-09-252-991A-25366 Sequence 25366, A
36	34	53.1	654	4	US-08-392-828C-2 Sequence 2, Appl 1
37	34	53.1	654	3	US-09-330-945-2 Sequence 2, Appl 1
38	34	53.1	1119	3	US-09-396-651B-2 Sequence 2, Appl 1
39	34	53.1	2409	6	5180808-2 Patent No. 5180808
40	33	51.6	278	4	US-09-252-991A-27917 Sequence 27917, A
41	33	51.6	280	4	US-09-053-374A-5 Sequence 5, Appl 1
42	33	51.6	281	4	US-09-053-374A-2 Sequence 2, Appl 1
43	33	51.6	386	4	US-09-328-352-7679 Sequence 7679, Ap
44	33	51.6	394	1	US-07-621-193A-5 Sequence 5, Appl 1
45	33	51.6	394	1	US-08-018-489C-5 Sequence 5, Appl 1

ALIGNMENTS

RESULT 1
Sequence 2, Application US/08824707
Patent No. 5919688
GENERAL INFORMATION:
APPLICANT: Ferrer, Pau
APPLICANT: Diers, Ivan
APPLICANT: Hedegaard, Lisbeth
APPLICANT: Halkier, Torben
APPLICANT: Azenjo, Juan
APPLICANT: Savva, Demitris
TITLE OF INVENTION: No. 5919688e1 enzyme with beta-1,3-glucanase activity
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: No. 59196880 No. 5919688dsk of No. 5919688th America, Inc;
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/824,707
FILING DATE: 14-April-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeta A 35,127
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 4290, 204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 306 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-824-707-2
Query Match 78.1%; Score 50; DB 2; Length 306;
Best Local Similarity 75.0%; Pred. No. 0.15;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
1 SGEIDIEITGN 12
|||||||:|:|

Db 180 SGEIDIMENVG 191

RESULT 2

US-09-159-106-2
; Sequence 2, Application US/09159106
; Patent No. 6284509
; GENERAL INFORMATION:
; APPLICANT: Ferrer, Pau
; APPLICANT: Diers, Ivan
; APPLICANT: Haikier, Torben
; APPLICANT: Hedegaard, Lisbeth
; TITLE OF INVENTION: An Enzyme With -1,3-Glucanase
; FILE REFERENCE: 4693.204-US
; CURRENT APPLICATION NUMBER: US/09/159,106
; CURRENT FILING DATE: 1998-09-23
; EARLIER APPLICATION NUMBER: 0427/96
; EARLIER FILING DATE: 1996-12-04
; EARLIER APPLICATION NUMBER: 0885/96
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: PCT/DK97/00160
; EARLIER FILING DATE: 1997-04-14
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Oerskovia xanthineolytica
US-09-159-106-2

Query Match

Best Local Similarity 68.8%; Score 44; DB 3; Length 263;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDIETIG 11
Db 116 SGEIDIMENVG 126

RESULT 3

US-09-159-106-13
; Sequence 13, Application US/09159106
; Patent No. 6284509
; GENERAL INFORMATION:
; APPLICANT: Ferrer, Pau
; APPLICANT: Diers, Ivan
; APPLICANT: Haikier, Torben
; APPLICANT: Hedegaard, Lisbeth
; TITLE OF INVENTION: An Enzyme With -1,3-Glucanase
; FILE REFERENCE: 4693.204-US
; CURRENT APPLICATION NUMBER: US/09/159,106
; CURRENT FILING DATE: 1998-09-23
; EARLIER APPLICATION NUMBER: 0427/96
; EARLIER FILING DATE: 1996-12-04
; EARLIER APPLICATION NUMBER: 0885/96
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: PCT/DK97/00160
; EARLIER FILING DATE: 1997-04-14
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Oerskovia xanthineolytica
US-09-159-106-13

Query Match 68.8%; Score 44; DB 3; Length 303;
Best Local Similarity 72.7%; Pred. No. 2;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDIETIG 11

Db 169 SGEIDIMENVG 179

RESULT 4

US-09-159-106-11
; Sequence 11, Application US/09159106
; Patent No. 6284509
; GENERAL INFORMATION:
; APPLICANT: Ferrer, Pau
; APPLICANT: Diers, Ivan
; APPLICANT: Haikier, Torben
; APPLICANT: Hedegaard, Lisbeth
; TITLE OF INVENTION: An Enzyme With -1,3-Glucanase
; FILE REFERENCE: 4693.204-US
; CURRENT APPLICATION NUMBER: US/09/159,106
; CURRENT FILING DATE: 1998-09-23
; EARLIER APPLICATION NUMBER: 0427/96
; EARLIER FILING DATE: 1996-12-04
; EARLIER APPLICATION NUMBER: 0885/96
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: PCT/DK97/00160
; EARLIER FILING DATE: 1997-04-14
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Oerskovia xanthineolytica
US-09-159-106-11

Query Match

Best Local Similarity 68.8%; Score 44; DB 3; Length 435;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDIETIG 11
Db 169 SGEIDIMENVG 179

RESULT 5

US-09-266-225D-12
; Sequence 12, Application US/09266225D
; Patent No. 6573364
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishan
; APPLICANT: Kingsmore, Stephen
; APPLICANT: Tcherenev, Velizar
; TITLE OF INVENTION: Isolation and Characterization of Hermansky-Pudlak Syndrome (HPS) Protein Complexes and HPS Protein-Interacting Proteins
; FILE REFERENCE: 15966-523
; CURRENT APPLICATION NUMBER: US/09/266,225D
; CURRENT FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 816
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-266-225D-12

Query Match 62.5%; Score 40; DB 4; Length 816;
Best Local Similarity 80.0%; Pred. No. 34;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 EIDIETIG 12
Db 54 EYDIETIG 63

RESULT 6

US-09-328-352-8132
; Sequence 8132, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328.352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8132
; LENGTH: 1380
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-8132

Query Match 62.5%; Score 40; DB 4; Length 1380;
Best Local Similarity 58.3%; Pred. No. 61;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 GEIDIETIGNR 13
Db 467 GEVDIDHGNR 478

RESULT 7
US-08-712-072C-4
; Sequence 4, Application US/08712072C
; Patent No. 5925541
; GENERAL INFORMATION:
; APPLICANT: Jack Goldstein, Alex Zhu and Lin Leng
; TITLE OF INVENTION: ENDO-BETA-GALACTOSIDASE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amster, Rothstein & Eberstein
; STREET: 90 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712.072C
; FILING DATE: 11-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Elizabeth A.
; REGISTRATION NUMBER: 39,911
; REFERENCE/DOCKET NUMBER: 63475/97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-5995
; TELEFAX: (212) 286-0854 or 286-0082
; TELER: TWX 710-581-4766
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 276 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGIN: SOURCE:
; ORGANISM: gub, Rhodothermus marinus

US-08-712-072C-4
Query Match 60.9%; Score 39; DB 2; Length 276;
Best Local Similarity 63.6%; Pred. No. 15;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SEIDIETIG 11
Db 146 NGEIDIMHWG 156

RESULT 8
US-08-737-526-4
; Sequence 4, Application US/08737526
; Patent No. 5871966
; GENERAL INFORMATION:
; APPLICANT: Kofoed, Lene Venke
; APPLICANT: Andersen, Lene No. 5871966boe
; APPLICANT: Kaupinen, Markus Sakari
; TITLE OF INVENTION: An Enzyme With Endo-1,3(4)-B-Glucanase
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NO. 58719660 No. 5871966disk of No. 5871966th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/737.526
; FILING DATE: 08-NOV-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Valeta, Gregg A.
; REGISTRATION NUMBER: 35,127
; REFERENCE/DOCKET NUMBER: 4174.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELER:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 292 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-737-526-4

Query Match 60.9%; Score 39; DB 2; Length 292;
Best Local Similarity 77.8%; Pred. No. 16;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEIDIETI 10
Db 158 GEIDIMETV 166

RESULT 9
US-09-098-580-4
; Sequence 4, Application US/09098580
; Patent No. 6140096
; GENERAL INFORMATION:
; APPLICANT: Kofoed, Lene Venke
; APPLICANT: Andersen, Lene No. 6140096boe
; APPLICANT: Kaupinen, Markus Sakari
; APPLICANT: Christgau, Stephan

;; TITLE OF INVENTION: An Enzyme With Endo-1,3(4)-B-Glucanase
;; TITLE OF INVENTION: Activity
;; NUMBER OF SEQUENCES: 6
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: No. 61400960 No. 6140096disk of No. 6140096th America, Inc.
;; STREET: 405 Lexington Avenue
;; CITY: New York
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10174-6401
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; OPERATING SYSTEM: IBM Compatible
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/098,580
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/737,526
;; FILING DATE: 08-NOV-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Valeta, Gregg A
;; REGISTRATION NUMBER: 35,127
;; REFERENCE/DOCKET NUMBER: 4174,204-US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-867-0123
;; TELEFAX: 212-878-9655
;;
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 292 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;;
US-09-098-580-4

Query Match 60.9%; Score 39; DB 3; Length 292;
Best Local Similarity 77.8%; Pred. No. 16;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEDIDITET 10
DB 158 GEDIDMETV 166

RESULT 10
US-09-252-991A-20691
; Sequence 20691, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20691
; LENGTH: 586
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20691

Query Match 60.9%; Score 39; DB 4; Length 586;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 3 EIDIDITIGN 12
DB 155 DIDLITGVN 164

RESULT 11
US-08-392-828C-37
; Sequence 37, Application US/08392828C
; Patent No. 5795962
; GENERAL INFORMATION:
; APPLICANT: IWANAGA, SADAHI
; APPLICANT: MUTA, TATSUSHI
; APPLICANT: SEKI, NORIAKI
; APPLICANT: ODA, TOSHIO
; TITLE OF INVENTION: NOVEL POLYPEPTIDE AND DNA ENCODING
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/392,828C
; FILING DATE: 28-FEB-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, PAULA A
; REGISTRATION NUMBER: 32,503
; REFERENCE/DOCKET NUMBER: FJN-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 262 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..262
; OTHER INFORMATION: /note="BGI A1 SEQUENCE (FIGURE 2)"
US-08-392-828C-37

Query Match 57.8%; Score 37; DB 1; Length 262;
Best Local Similarity 63.6%; Pred. No. 33;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGEIDITETG 11
DB 130 SGEIDVMEARG 140

RESULT 12
US-09-330-945-37
; Sequence 37, Application US/09330945
; Patent No. 6077946
; GENERAL INFORMATION:
; APPLICANT: IWANAGA, SADAHI
; APPLICANT: MUTA, TATSUSHI
; APPLICANT: SEKI, NORIAKI
; APPLICANT: ODA, TOSHIO
; TITLE OF INVENTION: DNA ENCODING HORSESHOE CRAB

TITLE OF INVENTION: AMBROCYTE LYSATE FACTOR G SUBUNIT A
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESS: PATENT ADMINISTRATOR, TESTA, HURWITZ &
ADDRESS: THIBEAULT, LLP
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/330,945
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/119,995
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: RIN-032DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 262 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein
LOCATION: 1..262
OTHER INFORMATION: /note="BGI A1 SEQUENCE (FIGURE 2)"
US-09-330-945-37
Query Match 57.8%; Score 37; DB 3; Length 262;
Best Local Similarity 63.6%; Pred. No. 33;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 SGEIDIIETIG 11
Db 130 SGEIDVMEARG 140
RESULT 13
US-08-712-072C-3
Sequence 3, Application US/08712072C
Patent No. 5925541
GENERAL INFORMATION:
APPLICANT: Jack Goldstein, Alex Zhu and Lin Leng
TITLE OF INVENTION: ENDO-BETA-GALACTOSIDASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESS: Amster, Rochstein & Ebenstein
STREET: 90 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE DISKETTE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/712,072C

FILING DATE: 11-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Elizabeth A.
REGISTRATION NUMBER: 39,911
REFERENCE/DOCKET NUMBER: 63475/97
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: el3D, Bacillus circulans
US-08-712-072C-3
Query Match 57.8%; Score 37; DB 2; Length 321;
Best Local Similarity 63.6%; Pred. No. 42;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 SGEIDIIETIG 11
Db 189 SGEIDVMEARG 199
RESULT 14
US-09-092-508-2
Sequence 2, Application US/09092508
Patent No. 6291643
GENERAL INFORMATION:
APPLICANT: Henzel, William J.
TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESS: Merchant, Gould, Smith, Edell, Welter & Schmidt
STREET: 3100 No. 6291643west Center, 90 South Seventh St
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/092,508
FILING DATE: 05-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/048,807
FILING DATE: 05-JUN-1997
APPLICATION NUMBER: 60/055,258
FILING DATE: 07-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ketelberger, Ph.D., Denise M
REGISTRATION NUMBER: 33,924
REFERENCE/DOCKET NUMBER: 11669.6USU1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-332-5300
TELEFAX: 612-332-9081
TELEX:

Search completed: September 16, 2003, 11:30:24
Job time : 2.37531 secs

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1194 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-09-092-508-2

Query Match 56.2%; Score 36; DB 3; Length 1194;
Best Local Similarity 46.2%; Pred. NO. 2.9e+02;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 SGEIDIIETIGNR 13
DB 928 NGAIEILLYVNR 940

RESULT 15
US-09-435-115-2
Sequence 2, Application US/09435115
Patent No. 6346607
GENERAL INFORMATION:
APPLICANT: Henzel, William J.
TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
STREET: 3100 No. 6346607 West Center, 90 South Seventh St
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/435,115
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/092,508
FILING DATE:
APPLICATION NUMBER: 60/055,258
FILING DATE: 07-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Kettelberger, Ph.D., Denise M
REGISTRATION NUMBER: 33,924
REFERENCE/DOCKET NUMBER: 11669.GUSU1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-332-5300
TELEFAX: 612-332-9081
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1194 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-09-435-115-2

Query Match 56.2%; Score 36; DB 4; Length 1194;
Best Local Similarity 46.2%; Pred. NO. 2.9e+02;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 SGEIDIIETIGNR 13
DB 928 NGAIEILLYVNR 940

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:22:42; Search time 1.63728 Seconds

(without alignments)
763.579 Million cell updates/sec

Title: US-09-596-101c-1

Perfect score: 64

Sequence: 1 SGEIDIIETIGNR 13

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR.76:*
2: PIR1:*
3: PIR2:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	70.3	294	2	D70525
2	45	70.3	499	2	JC6141
3	44	68.8	301	2	C87296
4	43	67.2	411	2	JC7869
5	43	67.2	1356	2	E87311
6	43	67.2	1377	2	A63345
7	43	67.2	1378	2	A62817
8	43	67.2	1411	2	C97595
9	41	64.1	736	2	T06757
10	40	62.5	815	2	B56708
11	40	62.5	1302	2	T43230
12	40	62.5	1342	1	RNECB
13	40	62.5	1342	1	S32680
14	40	62.5	1342	1	RNEBPT
15	40	62.5	1342	2	F91242
16	40	62.5	1342	2	A84934
17	40	62.5	1342	2	C86090
18	40	62.5	1342	2	AC0456
19	40	62.5	1342	1	AE0933
20	40	62.5	1342	1	H64073
21	40	62.5	1375	2	F82336
22	40	62.5	1389	2	T30824
23	40	62.5	1392	2	D82007
24	40	62.5	1394	2	A61236
25	40	62.5	1655	2	S47446
26	39	60.9	286	2	S48201
27	39	60.9	369	2	AH0451
28	39	60.9	642	2	B72428
29	39	60.9	665	2	S52072

30	39	60.9	795	2	A97627	hypothetical prote
31	39	60.9	795	2	AC2850	GEDF family prote
32	39	60.9	967	2	T48210	hypothetical prote
33	39	60.9	1324	2	T18265	endo-1,3(4)-beta-g
34	38	59.4	263	2	A72349	conserved hypotet
35	38	59.4	298	2	S75156	hypothetical prote
36	38	59.4	371	1	MMBCMK	inner membrane pro
37	38	59.4	371	2	F86096	inner membrane pro
38	38	59.4	371	2	B91256	inner membrane pro
39	38	59.4	544	2	T45498	ORF MSV140 hypothe
40	38	59.4	608	2	T28301	DNA-directed RNA p
41	38	59.4	1103	2	T06918	agarase (EC 3.2.1.
42	38	59.4	1143	2	S73173	hypothetical prote
43	37	57.8	309	1	EUSMAG	alkaline phosphata
44	37	57.8	316	2	T46000	
45	37	57.8	625	2	G97076	

ALIGNMENTS

RESULT 1

D70525
probable beta-1 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: D70525

R:Coile, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordo

Ralandream, M.A.; Rogers, J.; Devlin, K.; Peltell, T.; Gentles, S.; Hamlin, N.; Holroyd

Nature 393, 537-544, 1998

A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete gen

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: D70525

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-294 <COL>

A:Cross-references: GB:Z96800; GB:AL123456; NID:q3261800; PIDN:CAB09586.1; PID:e3216

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: RV0315

Query Match 70.3%; Score 45; DB 2; Length 294;

Best Local Similarity 75.0%; Pred. No. 1.6;

Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGEIDIIETIGN 12

DB 169 SGEIDIIETIGN 180

RESULT 2

JC6141
beta 1,3-glucanase (EC 3.2.1.-) precursor - sea urchin (Strongylocentrotus purpuratus

C:Species: Strongylocentrotus purpuratus (purple urchin)

C>Date: 11-Apr-1997 #sequence_revision 09-May-1997 #text_change 21-Jul-2000

C:Accession: JC6141; PC6037

R:Bachman, E.S.; McClay, D.R.

Proc. Natl. Acad. Sci. U.S.A. 93, 6808-6813, 1996

A>Title: Molecular cloning of the first metazoan beta-1,3 glucanase from eggs of the

A:Reference number: JC6141; MUID:96270625; PMID:8692900

A:Accession: JC6141

A:Molecule type: mRNA

A:Residues: 1-499 <BA2>

A:Cross-references: GB:U97711; NID:q1488256; PIDN:AAC47235.1; PID:q1488257

A:Accession: PC6037

A:Molecule type: protein

A:Residues: 21-40; 197-209; 329-344 <BA2>

A:Experimental source: egg

C:Comment: This enzyme functions in several extracellular activities including autoc

defense enzymes in plants.

C:Keywords: egg; glycosidase; hydrolase

F:1-20/Domain: signal sequence #status predicted <STG>
F:21-499/Product: beta 1,3-glucanase #status predicted <MAT>

Query Match 70.3%; Score 45; DB 2; Length 499;
Best Local Similarity 66.7%; Pred. No. 2.9;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDIIETIGN 12
|||||:|:|
Db 326 SGEIDLVESRGN 337

RESULT 3

C87296

beta-glucanase [Imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: C87296

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: C87296

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-301 <STO>

A:Cross-references: GB:AE005673; NID:q13421537; PIDN:AAK2367.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC0380

Query Match 68.8%; Score 44; DB 2; Length 301;
Best Local Similarity 80.0%; Pred. No. 2.5;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SGEIDIIETI 10
|||||:|:|
Db 152 SGEIDIIETI 161

RESULT 4

CJ7869

cell wall-lytic beta-1,3-glucan-hydrolase precursor, BglM - Bacillus circulans

C:Species: Bacillus circulans

C:Date: 09-Dec-2002 #sequence_revision 09-Dec-2002 #text_change 31-Mar-2003

C:Accession: JC7869; PC7199

R:Asano, T.; Takai, J.; Yamamoto, M.; Aono, R.

Biosci. Biotechnol. Biochem. 66, 1246-1255, 2002

A:Title: Cloning and structural analysis of bglM gene coding for the fungal cell wall-ly

A:Reference number: JC7869; MUID:22152179; PMID:12162545

A:Accession: JC7869

A:Molecule type: DNA

A:Residues: 1-411 <ASA>

A:Cross-references: DDBJ:AB078775

A:Accession: PC7199

A:Molecule type: protein

A:Residues: 28-42;95-111;152-170 <AS2>

C:Comment: This protein that is a fungal cell wall-degrading enzyme, hydrolyzes beta-1,3

C:Genetics:

A:Gene: bglM

Query Match 67.2%; Score 43; DB 2; Length 411;
Best Local Similarity 66.7%; Pred. No. 5.5;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SGEIDIIETIGN 12
|||||:|:|
Db 145 SGEIDIMERVNN 156

RESULT 5

E87311

DNA-directed RNA polymerase, beta subunit [Imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001

C:Accession: E87311

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: E87311

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1356 <STO>

A:Cross-references: GB:AE005673; NID:q13421681; PIDN:AAK2489.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC0502

Query Match 67.2%; Score 43; DB 2; Length 1356;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIIETIGNR 13
|||||:|:|
Db 448 GEIDIDIVLGNR 459

RESULT 6

AG3345

DNA-directed RNA polymerase (EC 2.7.7.6) [Imported] - Brucella melitensis (strain 16)

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 03-Jun-2002

C:Accession: AG3345

R:DelVecchio, V.G.; Kaprat, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov

; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Lei

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella meli

A:Reference number: AD3252; PMID:11756688

A:Accession: AG3345

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1377 <KUR>

A:Cross-references: GB:AE008917; PIDN:AAL51930.1; PID:q17982686; GSPDB:GN00190

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BME10749

A:Map position: 1

A:Superfamily: DNA-directed RNA polymerase beta chain

C:Keywords: nucleotidyltransferase

Query Match 67.2%; Score 43; DB 2; Length 1377;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIIETIGNR 13
|||||:|:|
Db 450 GEIDIDIVLGNR 461

RESULT 7

AB2817

DNA-directed RNA polymerase beta chain [Imported] - Agrobacterium tumefaciens (strain

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C:Accession: AB2817

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woc

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; MCC

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kar

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AB2817
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1378 <KUR>
A:Cross-references: GB:AE008688; PIDN:AAL42952.1; PID:g17740410; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: rpoB
A:Map position: circular chromosome
C:Superfamily: DNA-directed RNA polymerase beta chain

Query Match 67.2%; Score 43; DB 2; Length 1378;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GEIDIETIGNR 13
Db 450 GEIDIDINLGNR 461

RESULT 8
C97595
RNA polymerase beta chain (A0171070) [Imported] - Agrobacterium tumefaciens (strain C58,
C:Species: Agrobacterium tumefaciens
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C:Accession: C97595
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ouello, B.; Goldman,
A.; Liu, F.; Mollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2333-2328, 2001
A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: C97595
A:Molecule type: DNA
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1411 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK87716.1; PID:g15157080; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_3569
A:Map position: circular chromosome
C:Superfamily: DNA-directed RNA polymerase beta chain

Query Match 67.2%; Score 43; DB 2; Length 1411;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GEIDIETIGNR 13
Db 483 GEIDIDINLGNR 494

RESULT 9
T06757
hypothetical protein F15B8.180 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 22-Oct-1999
C:Accession: T06757
R:Queller, F.; Bens, V.; Rechmann, S.; Borikova, D.; Anserge, W.; Salenoubat, M.; Mewes,
submitted to the Protein Sequence Database, April 1999
A:Reference number: 215794
A:Accession: T06757
A:Molecule type: DNA
A:Residues: 1-736 <QDE>
A:Cross-references: EMBL:AL049660; GSPDB:GN00061; ATSP:F15B8.180
A:Experimental source: cultivar Columbia; BAC clone F15B8
C:Genetics:
A:Gene: ATSP:F15B8.180
A:Map position: 3
A:Intons: 114/3; 146/1; 208/2; 293/3; 365/3; 384/3; 429/3; 467/3; 536/2; 563/2; 640/3

Query Match 64.1%; Score 41; DB 2; Length 736;
Best Local Similarity 72.7%; Pred. No. 25;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 EIDIETIGNR 13
Db 109 EIDIETIGNR 119

RESULT 10
B56708
extracellular signal-regulated kinase 5 - human
C:Species: Homo sapiens (man)
C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 24-Sep-1999
C:Accession: B56708
R:Zhou, G.; Bao, Z.Q.; Dixon, J.E.
J. Biol. Chem. 270, 12665-12669, 1995
A>Title: Components of a new human protein kinase signal transduction pathway.
A:Reference number: A56708; MUID:95279403; PMID:7759517
A:Accession: B56708
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-815 <ZHO>
A:Cross-references: GB:U25278; NID:9837260; PIDN:AAA81381.1; PID:9837261
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase ho
C:Keywords: ATP
F:52-315/Domain: protein kinase homology <KIN>
F:60-68/Region: protein kinase ATP-binding motif

Query Match 62.5%; Score 40; DB 2; Length 815;
Best Local Similarity 80.0%; Pred. No. 43;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 EIDIETIGNR 12
Db 53 EVIDIETIGNR 62

RESULT 11
T43230
DNA-directed RNA polymerase (EC 2.7.7.6) beta chain - Spiroplasma citri
C:Species: Spiroplasma citri
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-May-2000
C:Accession: T43230
R:Laigret, F.; Gaurivaud, P.; Bove, J.
Gene 171, 95-98, 1996
A>Title: Unique organization of the rpoB region of Spiroplasma citri: a gene for a r
A:Reference number: 222353; MUID:96257200; PMID:8675039
A:Accession: T43230
A>Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-1302 <LA1>
A:Cross-references: EMBL:U25815; NID:9886051; PID:9886053; PIDN:AA044217.1
A:Experimental source: strain R8A2HP; specific_host Citrus sinensis
C:Genetics:
A:Gene: rpoB
A:Genetic code: SGC3
C:Superfamily: DNA-directed RNA polymerase beta chain
C:Keywords: nucleotidyltransferase; transcription

Query Match 62.5%; Score 40; DB 2; Length 1302;
Best Local Similarity 58.3%; Pred. No. 74;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GEIDIETIGNR 13
Db 475 GEVDIDHIGNR 486

RESULT 12
RNDB
DNA-directed RNA polymerase (EC 2.7.7.6) beta chain - Escherichia coli (strain K-12)
N:Alternate names: transcriptase beta chain
C:Species: Escherichia coli
C>Date: 31-Oct-1980 #sequence_revision 12-Dec-1997 #text_change 01-Mar-2002
C:Accession: F65205; A91109; JN0244; S12576; A91472; A00689; I52540; I52542; I52541;

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.
 .A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of *Escherichia coli* K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: F65205
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1342 <BLAT>
 A:Cross-references: GB:AE000472; GB:U00096; NID:g2367333; PIDN:AA23625.1; PID:g1790419;
 A:Experimental source: Strain K-12, substrain MG1655
 R;Ovchinnikov, Y.A.; Monastyrskaya, G.S.; Gubanov, V.V.; Guryev, S.O.; Chertov, O.Y.; MC
 Eur. J. Biochem. 116, 621-629, 1981
 A:Title: The primary structure of *Escherichia coli* RNA polymerase. Nucleotide sequence of
 A:Reference number: A91109; MUID:81260785; PMID:6266829
 A:Accession: A91109
 A:Molecule type: DNA
 A:Residues: 1-515, 'V', 517-1342 <OVCS>
 A:Cross-references: GB:V00339; GB:J01678; GB:K00449; NID:g42813; PIDN:CAA23625.1; PID:g4
 A:Note: most of this sequence was confirmed by amino acid analysis
 R;Ovchinnikov, Y.A.; Sverdlov, E.D.; Lipkin, V.M.; Monastyrskaya, G.S.; Chertov, O.Y.; G
 Bioorg. Khim. 6, 655-665, 1980
 A:Title: Primary structure of RNA polymerase from *E. coli*; nucleotide sequence of EcoRI-
 A:Reference number: JN0244
 A:Accession: JN0244
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 187-515, 'V', 517-1144 <OV2>
 R;Post, L.E.; Styrchak, G.D.; Nomura, M.; Lewis, H.; Dennis, P.P.
 Proc. Natl. Acad. Sci. U.S.A. 76, 1697-1701, 1979
 A:Title: Nucleotide sequence of the ribosomal protein gene cluster adjacent to the gene
 A:Reference number: S12572; MUID:79201667; PMID:377281
 A:Accession: S12572
 A:Molecule type: DNA
 A:Residues: 1-33 <POS>
 A:Cross-references: EMBL:V00339
 R;Delcuve, G.; Downing, W.; Lewis, H.; Dennis, P.P.
 Gene 11, 367-373, 1980
 A:Title: Nucleotide sequence of the proximal portion of the RNA polymerase beta subunit
 A:Reference number: A91472; MUID:81165543; PMID:7011900
 A:Accession: A91472
 A:Molecule type: DNA
 A:Residues: 1-105, 'G', 108-383, 'CSFRCSSP'
 A:Cross-references: GB:V00341; NID:g42823; PIDN:CAA3629.1; PID:g42824
 R;Gurevitch, A.I.; Avakov, A.E.; Kolosov, M.N.
 Bioorg. Khim. 5, 1735-1738, 1979
 A:Title: The nucleotide sequence at the proximal end of rpoB gene of *Escherichia coli*.
 A:Reference number: A00689
 A:Accession: A00689
 A:Status: significant sequence differences
 A:Molecule type: DNA
 R;Monastyrskaya, G.S.; Gubanov, V.V.; Guryev, S.O.; Lipkin, V.M.; Sverdlov, E.D.
 Bioorg. Khim. 6, 1423-1426, 1980
 A:Title: Primary structure of RNA polymerase from *E. coli*; Nucleotide sequence of the rpo
 A:Reference number: I52540
 A:Accession: I52540
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-188 <RES>
 A:Cross-references: GB:M38292; NID:g147717; PIDN:AA24579.1; PID:g147718
 R;Sverdlov, E.D.; Lipkin, V.M.; Monastyrskaya, G.S.; Gubanov, V.V.; Guryev, S.O.; Chert
 Bioorg. Khim. 6, 309-312, 1980
 A:Title: The nucleotide sequence of strong RNA polymerase binding site within the *E. coli*
 A:Reference number: I52542
 A:Accession: I52542
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 187-354 <RE2>
 A:Cross-references: GB:M38304; NID:g147719; PIDN:AA24580.1; PID:g551832
 R;Gurevich, A.I.; Igoshin, A.V.; Kolosov, M.N.
 Bioorg. Khim. 6, 1580-1584, 1980

A:Title: Structure of a central part of *E. coli* operon rpoBC. Nucleotide sequence of
 A:Reference number: I52541
 A:Accession: I52541
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1143-1342 <RE3>
 A:Cross-references: GB:M38303; NID:g147724; PIDN:AA24583.1; PID:g147725
 R;Monastyrskaya, G.S.; Gubanov, V.V.; Guryev, S.O.; Lipkin, V.M.; Sverdlov, E.D.
 Bioorg. Khim. 6, 1106-1109, 1980
 A:Title: Primary structure of EcoRI-F fragment of rpoB, C genes and corresponding fr
 A:Reference number: I52539
 A:Accession: I52539
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1143-1342 <RE4>
 A:Cross-references: GB:M38293; NID:g147721; PIDN:AA24581.1; PID:g147722
 C:Genetics:
 A:Gene: rpoB
 A:Map position: 90 min
 C:Complex: the active enzyme consists of the sigma chain and the core enzyme; the si
 released; the core enzyme is composed of two alpha chains, one beta chain, and one be
 C:Function:
 A:Description: part of the catalytic component of the active enzyme that catalyzes R
 C:Superfamily: DNA-directed RNA polymerase beta chain
 C:Keywords: nucleotidyltransferase; transcription

Query Match 62.5%; Score 40; DB 1; Length 1342;
 Best Local Similarity 58.3%; Pred. No. 76;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DB 440 GEVDIDHGNR 451

RESULT 13
 S32680
 DNA-directed RNA polymerase (EC 2.7.7.6) beta chain - *Buchnera aphidicola*
 C:Species: *Buchnera aphidicola*
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S32680
 R:Clark, M.A.; Baumann, L.; Baumann, P.
 submitted to the EMBL Data Library, April 1992
 A:Description: Sequence analysis of an aphid endosymbiont DNA fragment containing rpx
 A:Reference number: S32679
 A:Accession: S32680
 A:Molecule type: DNA
 A:Residues: 1-1342 <CLAB>
 A:Cross-references: EMBL:211913; NID:g296967; PIDN:CAA77970.1; PID:g296969
 C:Superfamily: DNA-directed RNA polymerase beta chain
 C:Keywords: nucleotidyltransferase; transcription

Query Match 62.5%; Score 40; DB 1; Length 1342;
 Best Local Similarity 58.3%; Pred. No. 76;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DB 440 GEVDIDHGNR 451

RESULT 14
 RNEBR
 DNA-directed RNA polymerase (EC 2.7.7.6) beta chain - *Salmonella typhimurium*
 C:Species: *Salmonella typhimurium*
 C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 11-Jun-1999
 C:Accession: S01794
 R:Lilistyn, N.A.; Monastyrskaya, G.S.; Sverdlov, E.D.
 Eur. J. Biochem. 177, 363-369, 1988
 A:Title: Genes coding for RNA polymerase beta subunit in bacteria. Structure/function
 A:Reference number: S01794; MUID:89052707; PMID:3036723
 A:Accession: S01794
 A:Molecule type: DNA

A:Residues: 1-1342 <LIS>
 A:Cross-references: GB:X04642; GB:M37431; GB:X04860; GB:X13854; NID:947918; PIDN:CAA2830
 C:Genetics:
 A:Gene: rpoB
 C:Superfamily: DNA-directed RNA polymerase beta chain
 C:Keywords: nucleotidyltransferase; transcription

Query Match 62.5%; Score 40; DB 1; Length 1342;
 Best Local Similarity 58.3%; Pred. No. 76;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 GEIDIIETIGNR 13
 ||:| |: |||
 Db 440 GEVDDIDHIGNR 451

RESULT 15

F91242
 RNA polymerase beta subunit [imported] - Escherichia coli (strain O157:H7, substrain RIM
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: F91242
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gasawara, N.; Yasunaga, T.; Kuwara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: F91242
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1342 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA038333.1; PID:Q13364386; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: ECs4910.
 C:Superfamily: DNA-directed RNA polymerase beta chain

Query Match 62.5%; Score 40; DB 2; Length 1342;
 Best Local Similarity 58.3%; Pred. No. 76;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 GEIDIIETIGNR 13
 ||:| |: |||
 Db 440 GEVDDIDHIGNR 451

Search completed: September 16, 2003, 11:29:36
 Job time : 3.63728 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:15:27 ; Search time 0.916877 Seconds

(without alignments)
666,771 Million cell updates/sec

Title: US-09-596-101c-1
Perfect score: 64
Sequence: 1 SEIDITITGNR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SWISSPROT_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	64.1	1343	1	RPOB_BUCBP
2	40	62.5	782	1	PRP_HSVT2
3	40	62.5	806	1	MR07_MOUSE
4	40	62.5	815	1	MR07_HUMAN
5	40	62.5	1302	1	RPOB_SPIC1
6	40	62.5	1342	1	RPOB_BUCAL
7	40	62.5	1342	1	RPOB_BUCAP
8	40	62.5	1342	1	RPOB_BUCOL
9	40	62.5	1342	1	RPOB_SALTY
10	40	62.5	1343	1	RPOB_HAETN
11	40	62.5	1375	1	RPOB_VIBCH
12	40	62.5	1392	1	RPOB_NEIMA
13	40	62.5	1392	1	RPOB_NEIMA
14	40	62.5	1655	1	N188_YEAST
15	39	60.9	286	1	GIB_RKMR
16	39	60.9	540	1	TCPO_CANAL
17	39	60.9	665	1	CNG_DROME
18	38	59.4	371	1	MAIK_ECOLI
19	38	59.4	1103	1	RPOB_CYPAD
20	38	59.4	1143	1	RPOB_PORPU
21	37	57.8	309	1	AKAR-STRO
22	37	57.8	682	1	EL3B_BACCI
23	37	57.8	1375	1	RPOB_COXBU
24	37	57.8	1492	1	AT7A_RAT
25	36	56.2	446	1	PGLX_COCOA
26	36	56.2	737	1	YMA2_YEAST
27	36	56.2	1096	1	RPOB_GUTHP
28	36	56.2	1131	1	RPOB_AMASP
29	36	56.2	1248	1	APAF_HUMAN
30	35.5	55.5	546	1	TCPO_SCHPO
31	35	54.7	57	1	Y160_ARCPO
32	35	54.7	151	1	VG14_BP22
33	35	54.7	261	1	Y100_MYCLE

34	35	54.7	313	1	K1PE_HAETN
35	35	54.7	393	1	DOM3_CAEEL
36	35	54.7	430	1	AROA_STAAV
37	35	54.7	432	1	AROA_STAAV
38	35	54.7	432	1	AROA_STAAV
39	35	54.7	462	1	YELL_RALSO
40	35	54.7	546	1	TRGD_ARCPC
41	35	54.7	690	1	CYG3_HUMAN
42	35	54.7	690	1	CYG3_RAT
43	35	54.7	691	1	CYG3_BOVIN
44	35	54.7	691	1	CYG3_MOUSE
45	35	54.7	728	1	ARPA_ECOLI

ALIGNMENTS

```

RESULT 1
RPOB_BUCBP
ID RPOB_BUCBP STANDARD: PRT: 1343 AA.
AC Q89B20;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE DNA-directed RNA polymerase beta chain (EC 2.7.7.6). (Transcriptase
DE beta chain) (RNA polymerase beta subunit).
GN RPOB OR BPO035.
OS Buchnera aphidicola (subsp. Balzongia pistaciae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=135842;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426901; PubMed=12522265;
RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,
RA Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,
RA Tamames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.;
RT "Reductive genome evolution in Buchnera aphidicola."
RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).
CC - FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC - CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC [RNA](n).
CC - SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
CC BETA' CHAIN.
CC - SIMILARITY: Belongs to the RNA polymerase beta chain family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: AE014016; AA026778.1;
DR PROSITE: PS01166; RNA_POL_BETA; 1.
KW Transference; Transcription; DNA-directed RNA polymerase;
KW Complete proteome.
SQ SEQUENCE 1343 AA; 151549 MW; 35A7FF57BC09F121 CRC64;
OY Query Match 64.1%; Score 41; DB 1; Length 1343;
OY Best Local Similarity 66.7%; Pred. No. 32;
OY Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Db 441 GEIDDIDHGNR 452

```

RESULT 2

```

ID PRTP_HSV12 STANDARD: PR1: 782 AA.
AC Q9WR6L: 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable processing and transport protein.
OS Herpesvirus tupaia (Strain 2) (TUV-2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae;
OX NCBI_TaxID=132678;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99319892; PubMed-10392721;
RA Bahr U., Springfield C., Tidona C.A., Darai G.;
RT "Structural organization of a conserved gene cluster of Tupaia
herpesvirus encoding the DNA polymerase, glycoprotein B, a probable
processing and transport protein, and the major DNA binding protein."
RL Virus Res. 60:133-136(1999).
CC -!- FUNCTION: THIS PROTEIN MAY AFFECT TRANSLOCATION OF THE VIRUS
GLYCOPROTEINS TO MEMBRANES. IT IS INVOLVED IN CAPSID MATURATION
(CC (BY SIMILARITY)).
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES PRTP FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF084543; AAD42934.1; -
DR InterPro; IPR000501; Proc_transpt.
DR Pfam; PF01366; PRTP; 1.
DR Capsid assembly.
SQ SEQUENCE 782 AA; 87392 MW; 4BE40314E5BAB9B CRC64;

Query Match 62.5%; Score 40; DB 1; Length 782;
Best Local Similarity 63.6%; Pred. No. 29;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 EIDIIETGNR 13
| | | | | | | |
DB 236 EQDIVEVGRS 246

RESULT 3
MK07_MOUSE STANDARD: PR1: 806 AA.
AC Q9WVS8: 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitogen-activated protein kinase 7 (EC 2.7.1.1.-) (extracellular signal-
DE regulated kinase 5) (ERK-5) (BMK1 kinase).
DE MAPK7 OR ERK5.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kamakura S., Moriyuchi T., Nishida E.;
RT "Activation of the protein kinase ERK5/BMK1 by receptor tyrosine
RT kinases: identification and characterization of a signaling pathway to
RT the nucleus."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MEK5 AND ERK5 INTERACT SPECIFICALLY WITH ONE ANOTHER AND
CC NOT WITH MEK1/ERK1 OR MEK2/ERK2 PATHWAYS (BY SIMILARITY).
CC -!- ENZYME REGULATION: Activated by tyrosine and threonine
CC phosphorylation (by similarity).

```

```

CC -1- DOMAIN: THE SECOND PROLINE-RICH REGION MAY INTERACT WITH ACTIN
CC TARGETING THE KINASE TO A SPECIFIC LOCATION IN THE CELL.
CC -1- PPM: AUTOPHOSPHORYLATED ON THREONINE AND TYROSINE RESIDUES. WHEN
CC THE C-TERMINAL PART OF THE KINASE, WHICH COULD HAVE A REGULATORY
CC ROLE, IS ABSENT (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/or\_send\_an\_email\_to\_license@1sb-sib.ch).
CC -----
CC DR EMBL; AB019373; BAA82039.1; -.
CC DR HSSP; P24941.1; HCL.
CC DR MGP; MG1:1346347; MapK7.
CC DR GO; GO:0045765; P:regulation of angiogenesis; IMP.
CC DR InterPro; IPR000719; ProL_kin.
CC DR InterPro; IPR000719; ProL_kin.
CC DR Pfam; PF00069; pkinase; 1.
CC DR ProDom; PD000001; ProL_kinase; 1.
CC DR SMART; SMO0220; S_TKc; 1.
CC DR PROSITE; PS01351; MAPK_1.
CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP_1.
CC DR PROSITE; PS00108; PROTEIN_KINASE_ST_1.
CC DR Transferrase; Ser/Thr/threonine-protein kinase; ATP-binding; Cell cycle;
CC phosphorylation.
CC KW DOMAIN 55 347 PROTEIN KINASE.
CC FT FT DOMAIN 434 465 PRO-RICH 1.
CC FT FT DOMAIN 521 524 POLY-ARG.
CC FT FT DOMAIN 578 700 PRO-RICH 2.
CC FT FT NP_BIND 61 69 ATP (BY SIMILARITY).
CC FT FT BINDING 84 84 ATP (BY SIMILARITY).
CC FT FT ACT_SITE 182 182 BY SIMILARITY.
CC FT FT MOD_RES 219 219 PHOSPHORYLATION (ACTIVATES THE KINASE)
CC FT FT MOD_RES 221 221 (BY SIMILARITY).
CC FT FT MOD_RES 221 221 PHOSPHORYLATION (ACTIVATES THE KINASE)
CC FT FT MOD_RES 221 221 (BY SIMILARITY).
CC FT FT MOD_RES 221 221 PHOSPHORYLATION (ACTIVATES THE KINASE)
CC FT FT MOD_RES 221 221 (BY SIMILARITY).
CC SQ SEQUENCE 806 AA; 87732 MW; E7CC41C4BBD0633 CRC64;
CC -----
CC Query Match 62.5%; Score 40; DB 1; Length 806;
CC Best Local Similarity 80.0%; Pred. No. 29;
CC Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
CC -----
CC QY 3 EDITION 12
CC 1:|||||||
CC 54 EYELETIGN 63
CC -----
CC RESULT 4
CC MK07_HUMAN STANDARD; PRT; 815 AA.
CC AC Q13164; Q16634;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Mitogen-activated protein kinase 7 (EC 2.7.1.1.-) (Extracellular signal-
CC regulated kinase 5) (ERK-5) (ERK4) (BMK1 kinase).
CC GN MAPK7 OR PRKM7 OR ERK5 OR ERK4.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
CC NX NCBI_TaxID=9606;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Fetal brain;
CC RX MEDLINE=95279403; PubMed=7759517;
CC ZHOU G., Bao Z.Q., Dixon J.E.;

```


RT "Components of a new human protein kinase signal transduction
RT pathway".
RT U. Biol. Chem. 270:12665-12669(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Placenta;
RX MEDLINE-95374539; PubMed-7646528;
RA Lee J.-D., Ulevitch R.J., Han J.;
RT "Primary structure of BMK1, a new mammalian map kinase";
RL Biochem. Biophys. Res. Commun. 213:715-724(1995).
CC -1- FUNCTION: MEK5 AND ERK5 INTERACT SPECIFICALLY WITH ONE ANOTHER AND
CC NOT WITH MEK1/ERK1 OR MEK2/ERK2 PATHWAYS.
CC -1- ENZYME REGULATION: Activated by tyrosine and threonine
CC phosphorylation (by similarity).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY ADULT TISSUES. ABUNDANT IN
CC HEART, PLACENTA, LUNG, KIDNEY AND SKELETAL MUSCLE. NOT DETECTABLE
CC IN LIVER.
CC -1- DOMAIN: THE SECOND PROLINE-RICH REGION MAY INTERACT WITH ACTIN
CC TARGETING THE KINASE TO A SPECIFIC LOCATION IN THE CELL.
CC -1- PTM: AUTOPHOSPHORYLATED ON THREONINE AND TYROSINE RESIDUES, WHEN
CC THE C-TERMINAL PART OF THE KINASE, WHICH COULD HAVE A REGULATORY
CC ROLE, IS ABSENT.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U25278; AAA81381.1; -
DR EMBL: U29725; AAA82931.1; -
DR EMBL: U29726; AAA82932.1; -
DR EMBL: U29727; AAA82933.1; -
DR PIR: B56708; B56708.
DR HSSP: P24941; IHCL.
DR Genew; HGNC:6880; MAPK7.
DR MIM; 602521; -
DR GO; GO:0004707; F:MAP kinase activity; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR001527; MAP_Kin.
DR InterPro; IPR007719; Prot_Kinase.
DR InterPro; IPR002290; Ser_Thr_Pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR Prodom; PD000001; Prot_Kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS01351; MAPK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KM Transferrase; Serine/threonine-protein kinase; ATP-binding; Cell cycle;
KM Phosphorylation.
FT DOMAIN 54 346 PROTEIN KINASE.
FT DOMAIN 337 340 POLY-ALA.
FT DOMAIN 433 464 PRO-RICH 1.
FT DOMAIN 520 523 POLY-ARG.
FT DOMAIN 577 699 PRO-RICH 2.
FT NP_BIND 60 68 ATP (BY SIMILARITY).
FT BINDING 83 83 ATP (BY SIMILARITY).
FT ACT_SITE 181 181 BY SIMILARITY.
FT MOD_RES 218 218 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT (BY SIMILARITY).
FT MOD_RES 220 220 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT (BY SIMILARITY).
FT CONFLICT 19 32 ARGGTRPHRLCS -> GPVKVEPAAASVA (IN
FT REF. 2).
FT CONFLICT 609 609 L -> V (IN REF. 2).
SQ SEQUENCE 815 AA; 88636 MW; 379AD69803207CCF CRC64;
Query Match 62.5%; Score 40; DB 1; Length 815;

Best Local Similarity 80.0%; Pred. No. 30;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 3 EIDIIETIGN 12
DB 53 EYEIIETIGN 62
RESULT 5
ID PROB_SPIC1 STANDARD; PRT; 1302 AA.
AC P47767;
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase
DE beta chain) (RNA polymerase beta subunit).
GN PROB.
OS Spiroplasma citri.
OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
OC Spiroplasmataceae; Spiroplasma.
OX NCBI_TaxID=2133;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RA82HP;
RX MEDLINE-96257200; PubMed-8675039;
RA Laigret F., Gaurivaud P., Bove J.;
RT "The unique organization of the rpoB region of Spiroplasma citri: a
RT restriction and modification system gene is adjacent to rpoB";
RL Gene 171:95-98(1996).
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC [RNA](N).
CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
CC BETA' CHAIN.
CC -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U25815; AAC44217.1; -
DR PIR: T43230; T43230.
DR HSSP; Q9KWD7; IHQM.
DR InterPro; IPR001572; RNA_POL_B.
DR Pfam; PF04563; RNA_POL_Rpb2_1; 1.
DR Pfam; PF04561; RNA_POL_Rpb2_2; 2.
DR Pfam; PF04565; RNA_POL_Rpb2_3; 1.
DR Pfam; PF00562; RNA_POL_Rpb2_6; 1.
DR Pfam; PF04560; RNA_POL_Rpb2_7; 1.
DR PROSITE; PS01166; RNA_POL_BETA; 1.
KM Transferrase; Transcription; DNA-directed RNA polymerase.
SQ SEQUENCE 1302 AA; 146533 MW; 8D971C23EB9FEB2F CRC64;
Query Match 62.5%; Score 40; DB 1; Length 1302;
Best Local Similarity 58.3%; Pred. No. 47;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
OY 2 GEIDIIETIGNR 13
DB 475 GEVDIDHIGNR 486
RESULT 6
RPOB_BUCAI
ID RPOB_BUCAI STANDARD; PRT; 1342 AA.

```

AC  P57146;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase
DE  beta chain) (RNA polymerase beta subunit).
GN  RPOB OR BU034.
OS  Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS  symbiotic bacterium).
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC  Enterobacteriaceae; Buchnera.
OX  NCBI_TaxID=118099;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  STRAIN=Tokyo 1998;
RA  MEDLINE=20445173; PubMed=10993077;
RA  Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT  "Genome sequence of the endocellular bacterial symbiont of aphids
RT  Buchnera sp. APS."
RL  Nature 407:81-86(2000).
CC  -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC  OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC  SUBSTRATES.
CC  -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC  (RNA)(N).
CC  -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
CC  ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
CC  BETA' CHAIN.
CC  -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: AP001118; BAB12761.1; -
DR  HSSP: Q9KMU7; IHOM.
DR  InterPro: IPR001572; RNA_POL_B.
DR  Pfam: PF04563; RNA_POL_Rpb2_1; 1.
DR  Pfam: PF04561; RNA_POL_Rpb2_2; 1.
DR  Pfam: PF04565; RNA_POL_Rpb2_3; 1.
DR  Pfam: PF00562; RNA_POL_Rpb2_6; 1.
DR  Pfam: PF04560; RNA_POL_Rpb2_7; 1.
DR  PROSITE: PS01166; RNA_POL_BETA; 1.
KW  Transference; Transcription; DNA-directed RNA polymerase;
KW  Complete proteome.
SQ  SEQUENCE 1342 AA; 151467 MW; 95252459873DF940 CRC64;

Query Match 62.5%; Score 40; DB 1; Length 1342;
Best Local Similarity 58.3%; Pred. No. 48;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY  2 GEIDIETIGNR 13
DB  440 GEVDIDIDHGNR 451

RESULT 7
ID  RPOB_BUCAP STANDARD; PRT; 1342 AA.
AC  P41184;
DT  01-FEB-1995 (Rel. 31, Created)
DT  01-FEB-1995 (Rel. 31, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase
DE  beta chain) (RNA polymerase beta subunit).
GN  RPOB OR BUSG035.
OS  Buchnera aphidicola (subsp. Schizaphis graminum).
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC  Enterobacteriaceae; Buchnera.

```

```

OX  NCBI_TaxID=98794;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=93160925; PubMed=1369199;
RT  Clark M.A., Baumann L., Baumann P.;
RT  "Sequence analysis of an aphid endosymbiont DNA fragment containing
RT  rpoB (beta-subunit of RNA polymerase) and portions of rplL and rpoC."
RL  Curr. Microbiol. 25:283-290(1992).
RN  [2]
RP  SEQUENCE FROM N.A.
RA  MEDLINE=22084549; PubMed=12089438;
RA  Tamas I., Klasson L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,
RA  Wernegreen J.J., Sandstroem J.P., Moran N.A., Andersson S.G.E.;
RT  "50 million years of genomic stasis in endosymbiotic bacteria."
RL  Science 296:2376-2379(2002).
CC  -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC  OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC  SUBSTRATES.
CC  -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC  (RNA)(N).
CC  -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
CC  ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
CC  BETA' CHAIN.
CC  -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: Z11913; CAAT7970.1; -
DR  EMBL: AE014079; AA67606.1; -
DR  PIR: S32680;
DR  HSSP: Q9KMU7; IHOM.
DR  InterPro: IPR001572; RNA_POL_B.
DR  Pfam: PF04563; RNA_POL_Rpb2_1; 1.
DR  Pfam: PF04561; RNA_POL_Rpb2_2; 1.
DR  Pfam: PF04565; RNA_POL_Rpb2_3; 1.
DR  Pfam: PF00562; RNA_POL_Rpb2_6; 1.
DR  Pfam: PF04560; RNA_POL_Rpb2_7; 1.
DR  PROSITE: PS01166; RNA_POL_BETA; 1.
KW  Transference; Transcription; DNA-directed RNA polymerase;
KW  Complete proteome.
SQ  SEQUENCE 1342 AA; 152060 MW; 7E30BC9E013D31ED CRC64;

Query Match 62.5%; Score 40; DB 1; Length 1342;
Best Local Similarity 58.3%; Pred. No. 48;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY  2 GEIDIETIGNR 13
DB  440 GEVDIDIDHGNR 451

RESULT 8
ID  RPOB_ECOLI STANDARD; PRT; 1342 AA.
AC  P00575; P00576; P78242;
DT  21-JUL-1986 (Rel. 01, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase
DE  beta chain) (RNA polymerase beta subunit).
GN  RPOB OR GRON OR NITB OR RIF OR RON OR STL OR STV OR TABD OR B3987 OR
GN  C4944 OR Z5560 OR ECS4910 OR SFA060.
OS  Escherichia coli.
OS  Escherichia coli O6.
OS  Escherichia coli O157:H7, and
OS  Shigella flexneri.
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

```

OC Enterobacteriaceae; Escherichia.
 OX NCBI_Taxid=562, 217992, 83334, 623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-E.coli;
 RX MEDLINE=81260785; PubMed=6266829;
 RA Ovechinnikov Y.A., Monastyrskaya G.S., Gubanov V.V., Guryev S.O.,
 RA Chertkov O.Y., Modyanov N.N., Grinkevich V.A., Makarova I.A.,
 RA Marchenko T.V., Polovnikova I.N., Lipkin V.M., Sverdlov E.D.;
 RT "The primary structure of Escherichia coli RNA polymerase. Nucleotide
 RT sequence of the rpoB gene and amino-acid sequence of the
 RT beta-subunit.";
 RL Eur. J. Biochem. 116:621-629(1981).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES-E.coli; STRAIN=12 / MG1655;
 RX MEDLINE=94089392; PubMed=8265357;
 RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
 RA Daniels D.L.;
 RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
 RT region from 89.2 to 92.8 minutes.";
 RL Nucleic Acids Res. 21:5408-5417(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES-E.coli; STRAIN=B;
 RA Miller E.S., Shih G.C., Chung S.K., Ballard D.N.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES-E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;
 RX MEDLINE=22388234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Rasko D., Buckles E.L., Liu S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 RT of uropathogenic Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES-E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirpatrick H.A.,
 RA Postif G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Groetbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamocis K.,
 RA Apodaca J., Anantharaman T.S., Llin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
 RL Nature 409:529-533(2001).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC SPECIES-E.coli; STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsudo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa K., Ogasawara N., Yasunaga T.,
 RA Kulara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 RN [7]
 RP SEQUENCE OF 1-33 FROM N.A.
 RC SPECIES-E.coli;
 RX MEDLINE=79201667; PubMed=377281;
 RA Post L.E., Strychartz G.D., Nomura M., Lewis H., Dennis P.P.;
 RT "Nucleotide sequence of the ribosomal protein gene cluster adjacent
 RT to the gene for RNA polymerase subunit beta in Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 76:1697-1701(1979).
 RN [8]
 RP SEQUENCE OF 1-391 FROM N.A.
 RC SPECIES-E.coli;
 RX MEDLINE=81165543; PubMed=7011900;

RA Delcuve G., Downing W., Lewis H., Dennis P.P.;
 RT "Nucleotide sequence of the proximal portion of the RNA polymerase
 RT beta subunit gene of Escherichia coli.";
 RL Gene 11:367-373(1980).
 RN [9]
 RP SEQUENCE OF 1-188 FROM N.A.
 RC SPECIES-E.coli;
 RA Gurevich A.I., Avakov A.E., Kolosov M.N.;
 RT "The nucleotide sequence at the proximal end of rpoB gene of
 RT Escherichia coli.";
 RL Bioorg. Khim. 5:1735-1739(1979).
 RN [10]
 RP SEQUENCE OF 1-188 FROM N.A.
 RC SPECIES-E.coli;
 RA Monastyrskaya G.S., Gubanov V.V., Guryev S.O., Lipkin V.M.,
 RA Sverdlov E.D.;
 RT "Primary structure of RNA polymerase from E.coli: Nucleotide sequence
 RT of the rpoB gene fragment and corresponding N-terminal amino acid
 RT sequence of the beta-subunit.";
 RL Bioorg. Khim. 6:1423-1426(1980).
 RN [11]
 RP SEQUENCE OF 187-354 FROM N.A.
 RC SPECIES-E.coli;
 RA Sverdlov E.D., Lipkin V.M., Monastyrskaya G.S., Gubanov V.V.,
 RA Guryev S.O., Chertkov O.Y.;
 RT "The nucleotide sequence of strong RNA polymerase binding site within
 RT the E.coli rpoB structural gene.";
 RL Bioorg. Khim. 6:309-312(1980).
 RN [12]
 RP SEQUENCE OF 187-1144 FROM N.A.
 RC SPECIES-E.coli;
 RA Ovechinnikov Y.A., Sverdlov E.D., Lipkin V.M., Monastyrskaya G.S.,
 RA Chertkov O.Y., Gubanov V.V., Guryev S.O., Modyanov N.N.,
 RA Grinkevich V.A., Makarova I.A., Marchenko T.V., Polovnikova I.N.;
 RT "Primary structure of RNA polymerase from E. coli: nucleotide sequence
 RT of EcoRI-C fragment of gene rpoB and amino acid sequence of the
 RT corresponding fragment of beta-subunit.";
 RL Bioorg. Khim. 6:655-665(1980).
 RN [13]
 RP SEQUENCE OF 1143-1342 FROM N.A.
 RC SPECIES-E.coli;
 RA Gurevich A.I., Igoshin A.V., Kolosov M.N.;
 RT "Structure of a central part of E.coli operon rpoBC. Nucleotide
 RT sequence of the gene for beta subunit of RNA polymerase.";
 RL Bioorg. Khim. 6:1580-1584(1980).
 RN [14]
 RP SEQUENCE OF 500-669 FROM N.A.
 RC SPECIES-E.coli; STRAIN=ATCC 25290;
 RA Wollet C., Drancourt M., Raoult D.;
 RT "RNA polymerase beta-subunit.";
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [15]
 RP SEQUENCE FROM N.A.
 RC SPECIES-S.flexneri; STRAIN=301 / Serotype 2a;
 RX MEDLINE=22272406; PubMed=12384590;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 RT through comparison with genomes of Escherichia coli K12 and O157.";
 RL Nucleic Acids Res. 30:4432-4441(2002).
 RN [16]
 RP MUTAGENESIS OF GLU-813.
 RC SPECIES-E.coli;
 RX MEDLINE=91296752; PubMed=2068078;
 RA Lee J., Kashley M., Borukhov S., Goldfarb A.;
 RT "A beta subunit mutation disrupting the catalytic function of
 RT Escherichia coli RNA polymerase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:6018-6022(1991).
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS

CC SUBSTRATES.
 CC CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
 CC (RNA)(N).
 CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
 CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
 CC BETA' CHAIN.
 CC -1- SIMILARITY: belongs to the RNA polymerase beta chain family.
 CC -1- CAUTION: REF.9 SEQUENCE DIFFERS FROM THAT SHOWN EXTENSIVELY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: V00339; CAA23625.1; -
 DR EMBL: V00340; CAA23627.1; -
 DR EMBL: U00006; AAC43085.1; -
 DR EMBL: AE000472; AAC76961.1; -
 DR EMBL: U76222; AAB18647.1; -
 DR EMBL: AE016770; AAN83372.1; ALT_INIT.
 DR EMBL: AE005630; AAC59183.1; -
 DR EMBL: AP002567; BAB38333.1; -
 DR EMBL: V00341; CAA23629.1; -
 DR EMBL: M38292; AAA24579.1; -
 DR EMBL: M38304; AAA24580.1; -
 DR EMBL: M38287; AAA24585.1; -
 DR EMBL: M38293; AAA24581.1; -
 DR EMBL: M38303; AAA24581.1; -
 DR EMBL: U77436; AAD09605.1; -
 DR EMBL: AE015413; AAN45489.1; ALT_INIT.
 DR PIR: F65205; RNECB.
 DR PIR: F91242; F91242.
 DR HSSP: Q9KMU7; IHOM.
 DR SWISS-2DPAGE: P00575; COLI.
 DR ECODBASE: D157.0; 67R EDITION.
 DR EcoGene: EG10894; rpoB.
 DR InterPro: IPR001572; RNA_pol_B.
 DR
 Query Match 62.5%; Score 40; DB 1; Length 1342;
 Best Local Similarity 58.3%; Pred. No. 48;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 Oy 2 GEIDDIETIGNR 13
 ||:|:|:|
 Db 440 GEVDDIDHIGNR 451.
 RESULT 9
 RPOB_SALTY STANDARD; PRT; 1342 AA.
 ID RPOB_SALTY STANDARD; PRT; 1342 AA.
 AC P06173; O9L9J8.
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase
 DE beta chain) (RNA polymerase beta subunit).
 GN RPOB OR STM4153 OR STMFL1.12.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sverdlov E.D., Litsyn N.A., Gurtev S.O., Monastyrskaya G.S.;
 RT Nucleotide sequence of the rpoB gene of Salmonella typhimurium coding
 RL for the beta-subunit of RNA polymerase.";
 RL Dokl. Biochem. 287:62-65(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89052707; PubMed=3056723;

RA Litsyn N.A., Monastyrskaya G.S., Sverdlov E.D.;
 RT "Genes coding for RNA polymerase beta subunit in bacteria.
 RT Structure/function analysis.";
 RL Eur. J. Biochem. 177:363-369(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=L72 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Potwilk S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Milvaney E.,
 RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 RT L72.";
 RL Nature 413:852-856(2001).
 RN [4]
 RP SEQUENCE OF 187-1144 FROM N.A.
 RX MEDLINE=86269165; PubMed=3015156;
 RA Sverdlov E.D., Litsyn N.A., Gurtev S.O., Smirnov Y.V.,
 RA Rostapshov V.M., Monastyrskaya G.S.;
 RT "Genes encoding the beta-subunit of bacterial RNA-polymerases. I.
 RT Primary structure of the EcoRI-C fragment of the Salmonella
 RT typhimurium gene rpoB.";
 RL Bioorg. Khim. 12:699-707(1986).
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 CC SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
 CC (RNA)(N).
 CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
 CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
 CC BETA' CHAIN.
 CC -1- SIMILARITY: belongs to the RNA polymerase beta chain family.
 CC
 CC this SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X04642; CAA28302.1; -
 DR EMBL: AF170176; AAF33499.1; -
 DR EMBL: AE008894; AAL22981.1; -
 DR EMBL: M38311; AAA27215.1; -
 DR PIR: S01794; RNEBET.
 DR HSSP: Q9KMU7; IHOM.
 DR styGene: SG10368; rpoB.
 DR InterPro: IPR001572; RNA_pol_B.
 DR Pfam: PF04563; RNA_pol_Rpb2_1; 1.
 DR Pfam: PF04561; RNA_pol_Rpb2_2; 1.
 DR Pfam: PF04565; RNA_pol_Rpb2_3; 1.
 DR Pfam: PF00562; RNA_pol_Rpb2_6; 1.
 DR Pfam: PF04560; RNA_pol_Rpb2_7; 1.
 DR PROSITE: PS01166; RNA_POL_BETA; 1.
 KW Transferrase; Transcription; DNA-directed RNA polymerase;
 KW Complete proteome.
 FT CONFLICT 401 401 G -> A (IN REF. 4).
 FT CONFLICT 542 542 R -> L (IN REF. 1, 2 AND 4).
 SQ SEQUENCE 1342 AA; 150600 MW; 8143964BFI709C02 CRC64;
 Oy 2 GEIDDIETIGNR 13
 ||:|:|:|
 Db 440 GEVDDIDHIGNR 451
 Query Match 62.5%; Score 40; DB 1; Length 1342;
 Best Local Similarity 58.3%; Pred. No. 48;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 RESULT 10

RPOB_HAEIN STANDARD: PRT: 1343 AA.

AC PA3738; 1343 AA.

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase beta chain) (RNA polymerase beta subunit).

GN RPOB OR H10515.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Haemophilus.

OX NCBI_Taxid=727;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Rd / KW20 / ATCC 51907;

RX MEDLINE=95350630; PubMed=7542800;

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M., McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D., Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M., Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D., Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C., Fine L.D., Fritchman J.L., Fuhrmann J.L., Geophagen N.S.M., RAE Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O., Venter J.C.;

RA "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.";

RL Science 269:496-512(1995).

CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS SUBSTRATES.

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA)(N).

CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1 BETA' CHAIN.

CC -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U3733; AAC22173.1; -

DR PIR: H64073; H64073.

DR HSSP: Q9KMW7; IHQM.

DR TIGR: H10515; -

DR InterPro: IPR001572; RNA_pol_B.

DR Pfam: PF04563; RNA_pol_Rpb2_1; 1.

DR Pfam: PF04561; RNA_pol_Rpb2_2; 1.

DR Pfam: PF04565; RNA_pol_Rpb2_3; 1.

DR Pfam: PF00562; RNA_pol_Rpb2_6; 1.

DR Pfam: PF04560; RNA_pol_Rpb2_7; 1.

DR PROSITE: PS01166; RNA_POL_BETA; 1.

KW Transferase; Transcription; DNA-directed RNA polymerase; Complete proteome.

KM SEQUENCE 1343 AA; 149783 MW; 4EF99CD648686444 CRC64;

SO

Query Match 62.5%; Score 40; DB 1; Length 1343;

Best Local Similarity 58.3%; Pred. No. 48;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIIETIGNR 13

Db 440 GEVDDIDHIGNR 451

RESULT 11

RPOB_VIBCH

ID RPOB_VIBCH STANDARD: PRT: 1375 AA.

AC Q9KV30; 1375 AA.

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase beta chain) (RNA polymerase beta subunit).

GN RPOB OR VC0328.

OS Vibrio cholerae.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;

OC Vibrionaceae; Vibrrio.

OX NCBI_Taxid=666;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=El Tor N16961 / Serotype O1;

RX MEDLINE=20406833; PubMed=10952301;

RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L., Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A., Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D., Ermolaeva M.D., Yamahayan J., Bass S., Qin H., Dragol I., Sellers P., McDonald L., Uterback T., Fleischmann R.D., Mierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C., Fraser C.M.;

RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.";

RL Nature 406:477-483(2000).

CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS SUBSTRATES (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA)(N).

CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1 BETA' CHAIN (BY SIMILARITY).

CC -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AE004121; AAF93501.1; -

DR PIR: F82336; F82336.

DR HSSP: Q9KMW7; IHQM.

DR TIGR: VC0328; -

DR InterPro: IPR001572; RNA_pol_B.

DR Pfam: PF04563; RNA_pol_Rpb2_1; 1.

DR Pfam: PF04561; RNA_pol_Rpb2_2; 1.

DR Pfam: PF04565; RNA_pol_Rpb2_3; 1.

DR Pfam: PF00562; RNA_pol_Rpb2_6; 1.

DR Pfam: PF04560; RNA_pol_Rpb2_7; 1.

DR PROSITE: PS01166; RNA_POL_BETA; 1.

KW Transferase; Transcription; DNA-directed RNA polymerase; Complete proteome.

KM SEQUENCE 1375 AA; 153285 MW; 1798976CD7225FDD CRC64;

SO

Query Match 62.5%; Score 40; DB 1; Length 1375;

Best Local Similarity 58.3%; Pred. No. 50;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIIETIGNR 13

Db 474 GEVDDIDHIGNR 485

RESULT 12

RPOB_NEIMA STANDARD: PRT: 1392 AA.

ID RPOB_NEIMA

AC P57009; 1392 AA.

DT 16-OCT-2001 (Rel. 40, Created)

Tue Sep 16 12:10:44 2003

us-09-596-101c-1.rsp

Page 10

Search completed: September 16, 2003, 11:26:28
job time : 1.91688 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:22:17 ; Search time 3.99496 Seconds

(without alignments)
839.729 Million cell updates/sec

Title: us-09-596-101c-1

Perfect score: 64

Sequence: 1 SGEIDIIETIGNR 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP rivirus:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	382	5	095VY3 lumbricus t
2	64	100.0	384	5	077072 eisenia foe
3	50	78.1	306	2	051333 oerskovia x
4	47	73.4	361	5	090064 pacifastacu
5	47	73.4	371	5	0879V2 aedes aegypt
6	46	71.9	366	5	08N0N3 penaeus mon
7	46	71.9	376	5	08MVS9 litopenaeus
8	45	70.3	294	16	007242 mycobacteri
9	45	70.3	462	3	09C236 neurospora
10	45	70.3	499	5	026660 strongyloce
11	44	68.8	301	16	09AB54 caulobacter
12	44	68.8	422	16	09EMR5 streptomyce
13	44	68.8	435	2	068641 oerskovia x
14	43	67.2	254	2	08GC27 lysobacter
15	43	67.2	278	3	096T05 alternaria
16	43	67.2	285	16	09F3A0 streptomyce

17	43	67.2	411	2	08KKH3	08kh3 bacillus ci
18	43	67.2	959	5	08SKF6	08sir6 encephalito
19	43	67.2	1356	16	09AAU2	09aau2 caulobacter
20	43	67.2	1377	16	08YHP8	08yhp8 brucella me
21	43	67.2	1377	16	08G069	08g069 brucella su
22	43	67.2	1378	16	09BN66	09bn66 rhizobium i
23	43	67.2	1380	16	0920H7	0920h7 rhizobium m
24	43	67.2	1383	2	09KJG4	09kjm5 bartonella
25	43	67.2	1383	2	09KJG5	09kjm5 bartonella
26	43	67.2	1411	16	08UE08	08ue08 agrobacteri
27	42	65.6	383	2	08GC25	08gc25 lysobacter
28	41	64.1	151	9	08H9G3	08h9g3 salmonella
29	41	64.1	736	10	09SVX7	09svx7 arabidopsis
30	40	62.5	82	11	08C4C6	08c4c6 mus musculu
31	40	62.5	342	2	09F404	09f404 neisseria m
32	40	62.5	342	2	09EU75	09eu75 neisseria m
33	40	62.5	342	2	09F405	09f405 neisseria m
34	40	62.5	342	2	09ET76	09et76 neisseria m
35	40	62.5	342	2	09F406	09f406 neisseria m
36	40	62.5	342	2	09ETJ6	09etj6 neisseria m
37	40	62.5	395	5	017492	017492 anopheles g
38	40	62.5	434	17	08TYW3	08tyw3 methanopyru
39	40	62.5	540	17	08TUC3	08tuc3 methanospir
40	40	62.5	575	16	08E3C3	08e3c3 streptococc
41	40	62.5	575	16	08DXQ4	08dxq4 streptococc
42	40	62.5	727	11	08R3I9	08r3i9 mus musculu
43	40	62.5	806	11	09WVP4	09wvp4 mus musculu
44	40	62.5	816	4	09GGS1	09ggs1 homo sapien
45	40	62.5	1304	5	08T126	08t126 dictyostell

ALIGNMENTS

RESULT 1	095VY3	PRELIMINARY:	PRT:	382 AA.
ID	095VY3			
AC	095VY3			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Coelomic cytolytic factor precursor.			
OS	Lumbricus terrestris (common earthworm).			
OC	Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;			
OC	Lumbricidae; Lumbricidae; Lumbricus.			
OX	NCBI_TaxID=6398;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Beschin A., De Baetselier P., Bilej M.;			
RT	"Distinct carbohydrate recognition domains of an earthworm defense			
RT	molecule recognize Gram negative and Gram positive bacteria."			
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF395805; AAL09587.1; "			
DR	InterPro: IPR000757; Glyco_hydro_16.			
DR	Pfam: PF00722; Glyco_hydro_16; 1.			
KW	Signal.			
FT	SIGNAL	1	17	POTENTIAL.
FT	SEQUENCE	382 AA;	43931 MW;	5256CF171EB7D3FB CRC64;
SO	Query Match	100.0%;	Score 64;	DB 5;
	Best Local Similarity	100.0%;	Pred. No. 0.0023;	Length 382;
	Matches 13;	Conservative 0;	Mismatches 0;	Indels 0;
			Gaps 0;	
QY	1 SGEIDIIETIGNR 13			
Db	178 SGEIDIIETIGNR 190			
RESULT 2	077072	PRELIMINARY:	PRT:	384 AA.
ID	077072			
AC	077072			
DT	01-NOV-1998 (TREMBLrel. 08, Created)			

DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE Coelomic cytolytic factor 1.
 GN CCF1.
 OS Euisenia foetida (Common brandling worm) (Common dung-worm).
 OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
 OC Lumbicina; Lumbiricidae; Euisenia.
 NCBI_TaxID=6396;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98406152; PubMed=9733802;
 RA Beschlin A., Billej M., Hanssens F., Raymakers J., Van Dyck E.,
 RA Revers H., Brys L., Gomez J., De Baetselier P., Timmermans M.;
 RT Identification and cloning of a glucan- and lipopolysaccharide-
 RT binding protein from Euisenia foetida earthworm involved in the
 RT activation of prophenoloxidase cascade."
 RL J. Biol. Chem. 273:24948-24954(1998).
 DR EMBL: AF030028; AAC35887.1;
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 SQ SEQUENCE 384 AA; 44322 MW; C90B5C94003BAD6D CRC64;

Query Match 100.0%; Score 64; DB 5; Length 384;
 Best Local Similarity 100.0%; Pred. No. 0.0023;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SGEIDIIETIGNR 13
 DB 178 SGEIDIIETIGNR 190

RESULT 3
 ID 051333 PRELIMINARY; PRT; 306 AA.
 AC 051333;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Beta-1,3-glucanase Ila.
 GN BGLITA
 OS Oerskovia xanthineolytica.
 OC Bacteria; Actinobacteriia; Actinobacteridae; Actinomycetales;
 OC Micrococciaceae; Promicromonosporaceae; Cellulosimicrobium.
 OX NCBI_TaxID=1710;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LL G109.
 RX MEDLINE=96345651; PubMed=8755914;
 RA Ferrer P., Halkier T., Hedegaard L., Savva D., Diers I., Azenjo J.A.;
 RT Nucleotide sequence of a beta-1,3-glucanase isoenzyme Ila gene of
 RT Oerskovia xanthineolytica LL G109 (Cellulomonas cellulans) and initial
 RT characterization of the recombinant enzyme expressed in Bacillus
 RT subtilis."
 RL J. Bacteriol. 178:4751-4757(1996).
 DR EMBL: U56935; AAC44371.1;
 DR HSSP: P23904; IAJK.
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 DR PROSITE: PS01034; GLYCOSTL_HYDROL_F16; 1.
 SQ SEQUENCE 306 AA; 32835 MW; CD8DB8C1A6F8DC04 CRC64;

Query Match 78.1%; Score 50; DB 2; Length 306;
 Best Local Similarity 75.0%; Pred. No. 0.7;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDIIETIGN 12
 DB 180 SGEIDIMENYGN 191

RESULT 4
 ID 09U0G4 PRELIMINARY; PRT; 361 AA.

AC 09U0G4;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE Lipopolysaccharide and beta-1,3-glucan binding protein precursor.
 GN LGBP.
 OS Pasteurella lenisusculus (Signal crayfish).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
 OC Astacoidae; Astacidae; Pacifastacus.
 OX NCBI_TaxID=6720;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hemocyte;
 RX MEDLINE=20092910; PubMed=10625682;
 RA Lee S., Wang R., Soderhall K.;
 RT "A lipopolysaccharide- and beta-1,3-glucan-binding protein from
 RT hemocytes of the freshwater crayfish Pacifastacus lenisusculus:
 RT purification, characterization, and cDNA cloning."
 RL J. Biol. Chem. 275:1337-1343(2000).
 DR EMBL: AJ250128; CAB65353.1;
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 KW Signal.
 FT SIGNAL
 FT CHAIN 1 15
 FT 16 361
 SQ SEQUENCE 361 AA; 41043 MW; 93A5E6791ED6619 CRC64;
 BINDING PROTEIN.

Query Match 73.4%; Score 47; DB 5; Length 361;
 Best Local Similarity 75.0%; Pred. No. 3;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDIIETIGN 12
 DB 172 SGEIDIVESRGN 183

RESULT 5
 ID 0879V2 PRELIMINARY; PRT; 371 AA.
 AC 0879V2;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Putative secreted protein.
 OS Aedes aegypti (Yellowfever mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 OC Neoptera; Megaloptera; Diptera; Nematocera; Culicoidae; Aedes.
 OX NCBI_TaxID=7159;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Black eye; TISSUE=Salivary gland;
 RA Valenzuela J.G., Pham V.M., Garfield M.K., Francischetti I.M.,
 RA Ribeiro J.M.C.;
 RT "Toward the stialome of the adult female mosquito Aedes aegypti."
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF466594; AAL76017.1;
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 SQ SEQUENCE 371 AA; 41943 MW; 1FE25DC7B665CE6 CRC64;

Query Match 73.4%; Score 47; DB 5; Length 371;
 Best Local Similarity 61.5%; Pred. No. 3.1;
 Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDIIETIGNR 13
 DB 199 SGEVDLMSRGNR 211

RESULT 6
 ID 08N0N3

```

ID 08NON3 PRELIMINARY; PRT: 366 AA.
AC 08NON3;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Beta-1,3-glucan binding protein.
OS Penaeus monodon (Penaeid shrimp).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
OC Penaeidae; Penaeus.
OX NCBI_TaxID=6687;
RN [1]
RP SEQUENCE FROM N.A.
RA Sittunyaluksana K., Lee S.Y., Soderhall K.;
RT "The beta-1,3-glucan binding protein from the black tiger shrimp,
RT Penaeus monodon."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL:AF368168; AAC21213.1;
DR InterPro: IPR000757; Glyco_hydro_16;
DR Pfam: PF00722; Glyco_hydro_16; 1.
SQ SEQUENCE 366 AA; 41497 MW; 6F540A60E83FDD7D CRC64;

Query Match
Best Local Similarity 71.9%; Score 46; DB 5; Length 366;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDILEITGN 12
DB 180 SGEIDILESRCN 191

RESULT 7
O8NVS9 PRELIMINARY; PRT: 376 AA.
AC 08NVS9;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Lipopolysaccharide and beta-1,3-glucan binding protein.
OS Litopenaeus stylirostris (blue shrimp).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
OC Penaeidae; Litopenaeus.
OX NCBI_TaxID=29019;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-Hepatopancreas;
RC MEDLINE=22068042; PubMed=12072514;
RX Roux M.W., Pain A., Kilmpel K.R., Dhar A.K.;
RT "The Lipopolysaccharide and beta-1,3-glucan Binding Protein Gene Is
RT Upregulated in White Spot Virus-Infected Shrimp (Penaeus
RT stylirostris)."
RL J. Virol. 76:7140-7149(2002).
DR EMBL:AF473579; AAM73871.1;
DR InterPro: IPR000757; Glyco_hydro_16;
DR Pfam: PF00722; Glyco_hydro_16; 1.
SQ SEQUENCE 376 AA; 42610 MW; D7ADC5A310208885 CRC64;

Query Match
Best Local Similarity 71.9%; Score 46; DB 5; Length 376;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDILEITGN 12
DB 190 SGEIDILESRCN 201

RESULT 8
O07242 PRELIMINARY; PRT: 294 AA.
ID 007242;
AC 007242;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)

```

```

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Probable beta-1,3-glucanase (Beta-glucanase, putative).
GN RV0315 OR MF0329 OR MTCY63.20.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=H37Rv;
RC MEDLINE=98295987; PubMed=9634230;
RX Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E., III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltham D., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandram M.A., Rogers J.,
RA Rutter S., Seeger K., Skellon S., Squares K.,
RA Sulston J.E., Taylor K., Whitehead S., Barrall B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=CDC 1551 / Oshkosh;
RC Fleisemann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Feldman J., Khoult H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z66800; CAB09586.1;
DR EMBL: AE006940; AAK44553.1;
DR TIGR: MF0329;
DR Tuberculist: Rv0315;
DR InterPro: IPR000757; Glyco_hydro_16;
DR InterPro: IPR006311; Tat.
DR Pfam: PF00722; Glyco_hydro_16; 1.
DR TIGRFAMs: TIGR01409; TAT_signal_seg. 1.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
KM Complete proteome.
FT CONFLICT 260 K -> T (IN REF. 2).
SQ SEQUENCE 294 AA; 32186 MW; 7B9897BA6740EBE0 CRC64;

Query Match
Best Local Similarity 70.3%; Score 45; DB 16; Length 294;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SGEIDILEITGN 12
DB 169 SGEIDILEWYGN 180

RESULT 9
O9C236 PRELIMINARY; PRT: 462 AA.
ID O9C236;
AC O9C236;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE Related to beta-1,3-glucan binding protein.
GN B7A16.110.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Algn V., Hohelsel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Meves H.W., Mannhaupt G.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

```

RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL513445; CAC28724.1; -
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 SO SEQUENCE 462 AA; 51868 MW; 28808B6DCAB3600 CRC64;

Query Match 70.3%; Score 45; DB 3; Length 462;
 Best Local Similarity 75.0%; Pred. No. 9.2;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGEIDIIETGN 12
 DB 269 SGEIDIMESRCN 280
 |||||:|:|:|

RESULT 10
 Q26660 PRELIMINARY; PRT; 499 AA.
 ID Q26660;
 AC Q26660;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Beta 1,3-glucanase.
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
 OC Strongylocentrotus.
 OX NCBI_TaxID=7668;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96270625; PubMed=8692900;
 RA Bachman E.S., McClay D.R.;
 RT "Molecular cloning of the first metazoan beta-1,3 glucanase from eggs
 of the sea urchin Strongylocentrotus purpuratus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:6808-6813(1996).
 DR EMBL: 049711; AAC47235.1;
 SO SEQUENCE 499 AA; 55275 MW; D863F36E0CF5AC CRC64;

Query Match 70.3%; Score 45; DB 5; Length 499;
 Best Local Similarity 66.7%; Pred. No. 10;
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGEIDIIETGN 12
 DB 326 SGEIDLVESRCN 337
 |||||:|:|:|

RESULT 11
 Q9AB54 PRELIMINARY; PRT; 301 AA.
 ID Q9AB54;
 AC Q9AB54;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Beta-glucanase.
 GN CC0380.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
 OC Caulobacteraceae; Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 Eissen J., Heldelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,

RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of *Caulobacter crescentus*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL: AB005711; NAK22367.1; -
 DR TIGR: CC0380;
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 KM Complete proteome.
 SO SEQUENCE 301 AA; 32526 MW; E15D6A603FF97823 CRC64;

Query Match 68.8%; Score 44; DB 16; Length 301;
 Best Local Similarity 80.0%; Pred. No. 8.8;
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEIDIIETI 10
 DB 152 SGEIDILETV 161
 |||||:|:|:|

RESULT 12
 Q9EWR5 PRELIMINARY; PRT; 422 AA.
 ID Q9EWR5;
 AC Q9EWR5;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Putative secreted hydrolase.
 DE SC00787 OR 3SCF60.19.
 GN Streptomyces coelicolor.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Seeger K.J., Harris D.;
 RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kleiser H.M., Denaparte D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 mb streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleiser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleiser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete *Streptomyces
 coelicolor* A3(2).";
 RL Nature 417:141-147(2002).
 DR EMBL: AL939106; CAC14352.1; -
 DR HSSP: P02879; 2AAT.
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR InterPro: IPR000722; Ricin_B_lectin.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 DR Pfam: PF00652; Ricin_B_lectin; 3.

DR SMART: SM00458; RICIN: 1.
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
 DR PROSITE: PS50231; RICIN_B_LLECTIN; 1.
 KW Hydrolase; Complete Proteome.
 SQ SEQUENCE 422 AA; 44143 MW; 7105047BE26F3020 CRC64;

Query Match 68.8%; Score 44; DB 16; Length 422;
 Best Local Similarity 72.7%; Pred. No. 13;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDIIETIG 11
 Db 170 SGEIDIMENYG 180

RESULT 13

O68641 PRELIMINARY; PRT; 435 AA.
 AC O68641;
 DT 01-AUG-1998 (TREMblrel. 07, Created)
 DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
 DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
 DE Beta-1,3-glucanase II.
 GN BGLII.
 OS Oerskovia xanthinolytica.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococciaceae; Promicromonosporaceae; Cellulosimicrobium.
 OX NCBI_TaxID=1710;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LL G109;
 RA Ventom A.M., Asenjo J.A.;
 RT "Characterization of yeast lytic enzymes from Oerskovia
 xanthinolytica LL-G109.";
 RL Enzyme Microb. Technol. 13:71-75(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LL G109;
 RX MEDLINE=96409238; PubMed=8814220;
 RA Parrado J., Escuredo P.R., Conejero-Lara F., Kotik M., Ponting C.P.,
 Asenjo J.A., Dobson C.M.;
 RT "Molecular characterization of a thermoactive beta-1,3-glucanase from
 Oerskovia xanthinolytica.";
 RL Biochim. Biophys. Acta 1296:145-151(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LL G109;
 RX MEDLINE=96250169; PubMed=8659924;
 RA Ferrer P., Hedegard L., Halkier T., Diers I., Savva D., Asenjo J.A.;
 RT "Molecular cloning of a lytic beta-1,3-glucanase gene from Oerskovia
 xanthinolytica LL G109.";
 RL Ann. N. Y. Acad. Sci. 782:555-566(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LL G109;
 RA Ferrer P., Andrews B.A., Asenjo J.A., Hedegard L., Diers I.;
 RT "BglII codes for a yeast-lytic beta-1,3-glucanase from Oerskovia
 xanthinolytica LL G109 (Cellulomonas cellulans) having a mannose-
 binding domain.";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF052745; AAC38290.1; -
 DR HSSP: P23904; IATK
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR InterPro: IPR000772; Ricin_B_lectin.
 DR Pfam: PF00652; Glyco_hydro_16; 1.
 DR SMART: SM00458; RICIN: 1.
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
 DR PROSITE: PS50231; RICIN_B_LLECTIN; 1.
 SQ SEQUENCE 435 AA; 46097 MW; 00F087BE644C0P58 CRC64;

Query Match 68.8%; Score 44; DB 2; Length 435;
 Best Local Similarity 72.7%; Pred. No. 13;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 OY 1 SGEIDIIETIG 11
 Db 169 SGEIDIMENYG 179

RESULT 14

O8GC27 PRELIMINARY; PRT; 254 AA.
 AC O8GC27;
 DT 01-MAR-2003 (TREMblrel. 23, Created)
 DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Beta-1,3-glucanase A.
 GN GLUA.
 OS Lysobacter enzymogenes.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Lysobacter.
 OX NCBI_TaxID=69;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-M4-7;
 RA Palumbo J.D., Sullivan R.F., Kobayashi D.Y.;
 RT "Molecular characterization of three beta-1,3-glucanase genes from
 Lysobacter enzymogenes strain M4-7 and partial biochemical
 RT characterization of their gene products by expression in Escherichia
 coli.";
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY157838; XANT7503.1; -
 SQ SEQUENCE 254 AA; 29131 MW; 0B80F73BA77D4A1B CRC64;

Query Match 67.2%; Score 43; DB 2; Length 254;
 Best Local Similarity 75.0%; Pred. No. 11;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SGEIDIIETIGN 12
 Db 135 SGEIDIMEHINN 146

RESULT 15

O96T05 PRELIMINARY; PRT; 278 AA.
 AC O96T05;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Mixed-linked glucanase (Fragment).
 OS Alternaria alternata.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
 OC Pleosporales; Pleosporaceae; mitosporic Pleosporaceae; Alternaria.
 OX NCBI_TaxID=5599;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Eshel D., Prusky D., Dinor A.;
 RT "Mixed-linked glucanase precursor of Alternaria alternata.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF282319; AAK69516.1; -
 DR InterPro: IPR000757; Glyco_hydro_16
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
 FT NON_TER 1
 FT NON_TER 278
 SQ SEQUENCE 278 AA; 31073 MW; B53AB749826B22D CRC64;

Query Match 67.2%; Score 43; DB 3; Length 278;
 Best Local Similarity 75.0%; Pred. No. 12;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SGEIDIIETIGN 12
 Db 128 SGEIDIAESRGN 139

Tue Sep 16 12:10:44 2003

us-09-596-101c-1.rspt

Page 6

Search completed: September 16, 2003, 11:28:40
Job time : 7.99496 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:14:52; Search time 117.038 Seconds
(without alignments)
520.781 Million cell updates/sec

Title: US-09-596-101c-3
Perfect score: 2240
Sequence: 1 MRWLVVLCILFEGGFATD.....DDEGDNAMQVYIRVYKRN 384

Scoring table:
BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: A_Geneseq_19Jun03.*

- 1: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
- 22: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
- 23: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
- 24: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2240	100.0	384	20	AA124914
2	357	15.9	492	22	AB862773
3	350	15.6	303	18	AAW29457
4	350	15.6	435	18	AAW29456
5	340.5	15.2	263	18	AAW29455
6	314	14.0	455	17	AAW29455
7	314	14.0	467	17	AAW29455
8	310	13.8	261	12	AAW29455
9	306	13.7	384	22	AAW29455

10	296	13.2	306	17	AAW29455
11	295	13.2	450	22	AB861180
12	286.5	12.8	422	19	AAW29455
13	249.5	11.1	410	22	AB860444
14	249.5	11.1	410	22	AB860451
15	209	9.3	673	16	AAW29455
16	208	9.3	673	16	AAW29455
17	184.5	8.2	845	22	AAW29455
18	168.5	7.5	341	20	AAW29455
19	138.5	6.2	294	22	AAW29455
20	131.5	5.9	292	17	AAW29455
21	121.5	5.4	424	19	AAW29455
22	115	5.1	975	23	AB860608
23	113.5	5.1	875	19	AAW29455
24	112.5	5.0	738	13	AAW29455
25	112.5	5.0	954	21	AAW29455
26	112.5	5.0	954	21	AAW29455
27	111.5	5.0	738	12	AAW29455
28	111	5.0	269	21	AAW29455
29	111	5.0	282	21	AAW29455
30	111	5.0	282	21	AAW29455
31	111	5.0	298	21	AAW29455
32	108	4.8	282	21	AAW29455
33	108	4.8	282	21	AAW29455
34	107.5	4.8	545	19	AAW29455
35	107	4.8	280	21	AAW29455
36	106	4.7	365	22	AAW29455
37	106	4.7	782	23	AAW29455
38	106	4.7	782	23	AAW29455
39	105.5	4.7	214	23	AAW29455
40	105	4.7	289	21	AAW29455
41	105	4.7	317	21	AAW29455
42	103.5	4.6	208	23	AAW29455
43	103.5	4.6	484	23	AAW29455
44	103	4.6	289	21	AAW29455
45	103	4.6	289	23	AAW29455

ALIGNMENTS

RESULT 1	AA124914	standard; Protein; 384 AA.
ID	AA124914	
XX	AA124914	
AC	AA124914	
XX	AA124914	
DT	25-AUG-1999	(first entry)
DE	Eisenia foetida coelomic cytolitic factor 1 protein.	
XX	Eisenia foetida; coelomic cytolitic factor 1; CCF-1; cancer;	
KW	trypanosomal infection; bacterial infection; tumour therapy;	
KW	inflammation; immunology.	
XX	Eisenia foetida.	
OS	Eisenia foetida.	
XX		
FT	Key	Location/Qualifiers
FT	Peptide	1..17
FT	Protein	18..384
FT		/label= CCF-1
XX		
PN	WO9931229-A2.	
PD	24-JUN-1999.	
XX		
PF	16-DEC-1998;	98WO-EP08169.
XX		
PR	17-DEC-1997;	97EP-0203974.
XX		
PA	(VLA-A) VLAAAS INTERUNIVERSITAIR INST BIOTECHNOG.	
XX		

Oerskovia beta-1,3
Drosophila melanog
Flavobacterium ker
Drosophila melanog
Drosophila melanog
(1-3)-beta-D-gluc
Clostridium perfr
Streptomyces sp. S
Mycobacterium tube
Tetrahymena harzia
Phaetia rhodozyma
Bifidobacterium lo
Bankia goudi endo
ADH complex protei
Ruminococcus flave
R. flavofaciens xyl
A. altocelligenes me
Arabidopsis thalia
Arabidopsis thalia
Herbicidally activ
Arabidopsis thalia
Arabidopsis thalia
Herbicidally activ
Gene 21 human secr
Human metabolism-a
Haemophilus influe
Bacterial lichenas
Arabidopsis thalia
Arabidopsis thalia
Bacterial lichenas
Listeria monocytog
Arabidopsis thalia
Herbicidally activ

PI De Baetselier P;
 XX WPI, 1999-365905/32.
 DR N-PSDB; AAX83611.
 XX
 PT Eissenla foetida polypeptides derived from coelomic cytolytic factor 1
 XX
 PS Claim 2; Page 48-49; 49pp; English.
 XX
 CC The present sequence represents the Eissenla foetida coelomic cytolytic
 CC factor 1 (CCF-1). The protein has antiparasitic, antibacterial and
 CC antiinflammatory activity. Recombinant coelomic cytolytic factor 1
 CC (rCCF-1) is trypanolytic for the African trypanosome Trypanosoma brucei
 CC in a dose-dependent manner. The trypanolytic activity of rCCF-1 can be
 CC inhibited by anti-CCF-1 and anti-tumour necrosis factor (TNF)/TIP
 CC monoclonal antibodies. Furthermore, N,N'-diacetylchitobiose inhibits
 CC potentially trypanolytic activity of rCCF-1. These data corroborate the
 CC findings that CCF-1 shares a trypanolytic, lectin-like domain with
 CC TNF-alpha. CCF-1 is useful to treat trypanosomal or bacterial infections
 CC or cancer. The proteins and peptides are also useful in tumour therapy,
 CC inflammation and other areas of immunology. The annexed peptide is
 CC derived from a 42 kDa cytolytic protein named CCF-1 that binds
 CC lipopolysaccharide and beta-1,3-glucan. The factor resembles the
 CC vertebrate tumour necrosis factor-alpha (TNF-alpha), and may be used as
 CC an alternative for TNF-alpha.
 XX
 SQ Sequence 384 AA;

Query Match 100.0%; Score 2240; DB 20; Length 384;
 Best Local Similarity 100.0%; Pred. No. 3.7e-202;
 Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRWTLVYLCILFGEFATMDQYHIVODEFDYFDGAKQHEVTATGSGSEFQLYTOD 60
 DB 1 MRWTLVYLCILFGEFATMDQYHIVODEFDYFDGAKQHEVTATGSGSEFQLYTOD 60
 QY 61 GANSFVRDGLFKPRLANINPOTGAPGTDFMNGVLDVWAMGACTNTDNGCYRT 120
 DB 61 GANSFVRDGLFKPRLANINPOTGAPGTDFMNGVLDVWAMGACTNTDNGCYRT 120
 QY 121 GAAGDIPRAPSARVRFQKSTFGHRYVYHAKMPVGMILMPALIMLPEDWYGGWPSGE 180
 DB 121 GAAGDIPRAPSARVRFQKSTFGHRYVYHAKMPVGMILMPALIMLPEDWYGGWPSGE 180
 QY 181 IDIIETIGNDFKNTGGEFLGIQKMGSTMHWGPMCDNRWLTSLPKHSDMNYGDNFHT 240
 DB 181 IDIIETIGNDFKNTGGEFLGIQKMGSTMHWGPMCDNRWLTSLPKHSDMNYGDNFHT 240
 QY 241 FWFDMSPNGLRFEVDENQALLDVPYPLIDANPMWVDFWEMGKFWLPQYENDNFWAGGTN 300
 DB 241 FWFDMSPNGLRFEVDENQALLDVPYPLIDANPMWVDFWEMGKFWLPQYENDNFWAGGTN 300
 QY 301 LAFPDNFHILNVAAGTGTFIPDGINRGDPALOKPWSNGDWYNDAMRKFPDARGNW 360
 DB 301 LAFPDNFHILNVAAGTGTFIPDGINRGDPALOKPWSNGDWYNDAMRKFPDARGNW 360
 QY 361 KMTWDEGDNNAQVDYIRYKRN 384
 DB 361 KMTWDEGDNNAQVDYIRYKRN 384

RESULT 2
 ABB62773
 ID ABB62773 standard; Protein; 492 AA.
 XX
 AC ABB62773;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 15111.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 pharmaceutical.

XX OS Drosophila melanogaster.
 XX
 XX PN W0200171042-A2.
 XX
 XX PD 27-SEP-2001.
 XX
 XX PF 23-MAR-2001; 2001WO-US09231.
 XX
 XX PR 23-MAR-2000; 2000US-191637P.
 XX
 XX PA 11-JUL-2000; 2000US-0614150.
 XX
 XX PE (PEKE) PE CORP NY.
 XX
 XX PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 XX WPI: 2001-656860/75.
 XX
 XX N-PSDB; ABL06876.
 XX
 XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
 XX genes from Drosophila and for elucidating cell signalling and cell-cell
 XX interactions -
 XX
 XX PS Disclosure: SEQ ID NO 15111; 21pp + Sequence Listing; English.
 XX
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 492 AA;

Query Match 15.9%; Score 357; DB 22; Length 492;
 Best Local Similarity 27.2%; Pred. No. 6.8e-25;
 Matches 101; Conservative 67; Mismatches 133; Indels 70; Gaps 16;

QY 26 IVODEFDYFDGAKQHEV-TATGSGSEFQLYTQDGANSPVROGKLFKPLLDNINP 84
 DB 175 LLEFETDQINESIMHIVRLPLDSKDAEFLY--DG-KAVYHDCNLTIEP-LTMSYRP 230
 QY 85 QTGAPGTDFMNGVLDVWAMYGACTNTDN--NGCY----RTGAAGDIPRAPSARVRFQ 138
 DB 231 DLS-----TANSLRD---LSEKCTGHNRKIKECILHSTSGSGGIMPPIVTPRISTKE 280
 QY 139 KYSFTGHGVYVYHAKMPVGMILMPALIMLP-EDWYGGWPSRGEIDIIETIGNRDKKNTGG 197
 DB 281 TFAFOYGKIEIRALPLPGDWITVPLILLEPTEWYGQSGEGQVALALAGSVLRMRG 340
 QY 198 EFLGIQKMGSTMHWGPMCDNRWLTSLPKHSDM-----NYGDNHFTFWFDMSPNGL 250
 DB 341 KLVD-----GRSLYGP-----VLSIDAHQEDLMLSKRKISHREDDEHTTSLDSSNRL 390
 QY 251 REFVDENQALLDVPYPLIDANPMWVDFWEMGKFWLPQYENDNFWAGGTNLAFPDONHF 310
 DB 391 LFSVQGVYGEMLNGFTELDENP-----RWKGGPMAFPDKMYI 430
 QY 311 ILNVAAGTGTFIPDGINRGDPALOKPWSNGDWYNDAMRKFPDARGNWKTWDEGDN 370
 DB 431 SLGVSVGFGFDF-----VDHLRTATYERKPAN--YHQAKLQFHQADQDOWLPTKOP--- 480
 QY 371 NAMOVYIRYV 381
 DB 481 -ALKIDIVRVF 490

RESULT 3


```

AAW29457
ID AAW29457 standard; Protein; 303 AA.
XX
AC AAW29457;
XX
DT 14-APR-1998 (first entry)
XX
DE Oerskovia xanthineolytica beta-1,3-glucanase.
XX
KM Beta-1,3-glucanase; lytic enzyme; yeast; beta glucan degradation;
  fungal cell wall; intracellular product; purification; protoplast.
XX
OS Oerskovia xanthineolytica LLG109 (DSM 10297).
XX
FH Key
FT Peptide 1..52
FT /label= Stg_peptide
FT Protein 53..303
FT /label= Mat_protein
XX
PN W09739114-A1.
XX
PD 23-OCT-1997.
XX
PF 14-APR-1997; 97WO-DK00160.
XX
PR 23-AUG-1996; 96DK-0000885.
PR 12-APR-1996; 96DK-0000427.
XX
PA (NOVO ) NOVO-NORDISK AS.
XX
PI Diers I, Ferrer P, Halkier T, Hedegaard L;
XX
DR WPI: 1997-526451/48.
DR N-PSDB: AAT89157.
XX
PT New isolated beta-1,3-glucanase enzyme - obtained from Oerskovia
PT xanthineolytica, used particularly for the lysis of microbial cells
PT for obtaining desirable products
XX
PS Example 2; Page 42-43; 64pp; English.
XX
CC This sequence comprises the polypeptide precursor of a novel
CC Oerskovia xanthineolytica enzyme that exhibits beta-1,3-glucanase
CC (Bg) activity. Its amino acid sequence was deduced from an
CC isolated genomic DNA sequence (see AAT89157). Claimed DNA constructs
CC that encode the novel Bg (see also AAW29455), a mannose binding
CC domain (see AAW29456) or a full-length enzyme, i.e. Bg with mannose
CC binding domain (see AAW29456), can be used to produce recombinant Bg
CC polypeptides, with or without a mannose binding domain, in fungal
CC or bacterial host cells. Bg polypeptides are used for the
CC degradation or modification of beta-glucan containing material,
CC especially for the gentle lysis of microbial cell walls, thereby
CC enabling recovery of desirable intracellular products with a
CC reduced amount of contaminants. They can also be used for the
CC production of e.g. pigments, colourants, flavourants, yeast
CC extracts, pharmaceuticals, food or feed compositions, and to
CC prepare protoplasts for use in fusion, transformation and cloning
CC studies.
XX
SQ Sequence 303 AA;
XX
Query Match 15.6%; Score 350; DB 18; Length 303;
Best Local Similarity 28.5%; Pred. No. 1.6e-24;
Matches 103; Conservative 30; Mismatches 96; Indels 132; Gaps 12;
XX
OY 26 IVMODEDYFDG-----AKMOHEVTATGGNSEFQVLTQDANSEFVRGKFIPTLLAD 80
DB 57 LMSDEPDGAGAGSPNPAVMNHETGAGHGMNAELQNTASRANSAL-DG----- 105
OY 81 NINPOTGAPFGCTDMYNGVLDMYAGACTNTDNNGCYRTGAADIPAMGARYTFOKY 140
DB 106 -----GNLVITA-----RREGDGS-----TSARKTTQCKY 131

```

```

OY 141 SEFHGRVYVHAKMPVDMWMLPAIMLPEDMVYGGWPRSGEIDIEIIGNRDFKNTGGEFL 200
DB 132 QPOYGRIFEARIDIPRQGGIMPAFMMLGSGSPGTWPWSSGEIDIMENVGFEPHVVHG---- 187
OY 201 GIQKMGSTHMGKPGWDDNRYWLTSLPKHSDDNRYGNFHTFWFDKSPNGIRFVVDENQA 260
DB 188 -----TVH-GPGVSGGS-GITGMVQHPOGWSFADTFHTFAVDMKPGELTFWVD--GQQ 236
OY 261 LLDVPEPLTDAMPWVDFWENWCKPMLPQYENDNPMAGSTNLAPFDGNFHLINVAVGSTN 320
DB 237 FHRVTRASYGANW-----VFDPFFLLILVAVGGQW 268
OY 321 GFIPDGCINRGDPLQKFPWSNGDWNDAARRKFFDARGKWKWTWDEGDNMAQVDYIRV 380
DB 269 PGYDGG-----TTQLP-----QQMKVDYVRV 289
OY 381 Y 381
DB 290 Y 290
XX
RESULT 4
AAW29456
ID AAW29456 standard; Protein; 435 AA.
XX
AC AAW29456;
XX
DT 14-APR-1998 (first entry)
XX
DE Oerskovia xanthineolytica beta-1,3-glucanase.
XX
KM Beta-1,3-glucanase; lytic enzyme; yeast; beta glucan degradation;
  fungal cell wall; intracellular product; purification; protoplast.
XX
OS Oerskovia xanthineolytica LLG109 (DSM 10297).
XX
FH Key
FT Peptide 1..52
FT /label= Stg_peptide
FT Protein 53..435
FT /label= Mat_protein
FT Domain 304..435
FT /label= Mannose-binding_domain
XX
PN W09739114-A1.
XX
PD 23-OCT-1997.
XX
PF 14-APR-1997; 97WO-DK00160.
XX
PR 23-AUG-1996; 96DK-0000885.
PR 12-APR-1996; 96DK-0000427.
XX
PA (NOVO ) NOVO-NORDISK AS.
XX
PI Diers I, Ferrer P, Halkier T, Hedegaard L;
XX
DR WPI: 1997-526451/48.
DR N-PSDB: AAT89156.
XX
PT New isolated beta-1,3-glucanase enzyme - obtained from Oerskovia
PT xanthineolytica, used particularly for the lysis of microbial cells
PT for obtaining desirable products
XX
PS Example 2; Page 39-40; 64pp; English.
XX
CC This sequence comprises the polypeptide precursor of a novel
CC Oerskovia xanthineolytica enzyme that exhibits beta-1,3-glucanase
CC (Bg) activity and which includes a mannose binding domain (MBD).
CC Its amino acid sequence was deduced from an isolated genomic DNA
CC sequence (see AAT89156). Claimed DNA constructs that encode the
CC novel Bg lacking a MBD (see AAW29455 and AAW29457), a MBD (see

```

CC AAM29456), or the full-length enzyme can be used to produce recombinant
 CC BG polypeptides, with or without a mannose binding domain, in fungal
 CC or bacterial host cells. BG polypeptides are used for the
 CC degradation or modification of beta-glucan containing material,
 CC especially for the gentle lysis of microbial cell walls, thereby
 CC enabling recovery of desirable intracellular products with a
 CC reduced amount of contaminants. They can also be used for the
 CC production of e.g. pigments, colourants, flavourants, yeast
 CC extracts, pharmaceuticals, food or feed compositions, and to
 CC prepare protoplasts for use in fusion, transformation and cloning
 CC studies.

CC Sequence 435 AA;

Query Match 15.6%; Score 350; DB 18; Length 435;
 Best Local Similarity 28.5%; Pred. No. 2,6e-24;
 Matches 103; Conservative 30; Mismatches 96; Indels 132; Gaps 12;

QY 26 IWMODEFDYFDG-----AKMOHEVTATGGNSEFOLYTODGANSFVRDGLFKFTLLAD 80
 DB 57 LMSDEPFGAAGSAPNPAVMNHETGAHGMNAELQNTTASRANSAL-DGQ----- 105
 QY 81 NINPQTGAPEFGTDFMYNGVLDMVAMYGACTNTDNGCYRTGAAGDIPPMASARVTFQKY 140
 DB 106 -----GNLVITA-----RREGDSY-----TSARMTTQGY 131
 QY 141 SFTHGRVYVAAKMPVGMVLPALWMLPEDWYTGMPRSGEIDIIETIGNRPFKNTGGEFL 200
 DB 132 QPOYGTIEARIQIPROGQIMPAFWMLGSGFPPTPSSGSDIDIMENVGFEPRHVG---- 187
 QY 201 GIOKMGSTMHGPGMDNRWLTSLPKHSDMANYGDNFHFPMWDSNGRLEFVNDENQA 260
 DB 188 -----TVH-GEYSGGS-GITGMTOHPQGSFADTFHFPAVDMKPELTFWVD--GQO 236
 QY 261 LLDVPEPLIDANPMWVDFEWEMGKFWLPQYENDNPMAGTNLAEFDQNFHFLNVAVGSTN 320
 DB 237 FHRVTRASVGANAM-----VFDDPEFLILNVAVGQW 268
 QY 321 GFIPDGCINRGDPALOKPMNSGNDWYNDAMRKFFDARGNMKWTVWDESDNNAQVDYIRV 380
 DB 269 PGYPDG-----TTQLP-----OQMKVDYVRV 289
 QY 381 Y 381
 DB 290 Y 290

RESULT 5
 AAM29455
 ID AAM29455 standard; Protein; 263 AA.

XX AAM29455;
 XX 14-APR-1998 (first entry)
 DE Oerskovia xanthineolytica mature beta-1,3-glucanase.
 XX
 XX Beta-1,3-glucanase; lytic enzyme; yeast; beta glucan degradation;
 KM fungal cell wall; intracellular product; purification; protoplast.
 XX
 OS Oerskovia xanthineolytica ILG109 (DSM 10297).

XX
 FH Key Location/Qualifiers
 FT CDS 23..955
 FT sig_peptide /tag- a
 FT /tag- b
 FT mat_peptide 164..952
 FT /tag- c
 XX
 PN WO9739114-A1.
 XX 23-OCT-1997.

XX
 PE 14-APR-1997; 97WC-DK00160.
 XX
 PR 23-AUG-1996; 96DK-0000885.
 PR 12-APR-1996; 96DK-0000427.
 XX
 PA (NOVO) NOVO-NORDISK AS.
 XX
 PI Diers I, Ferrer P, Halkier T, Hedegaard L;
 XX
 DR WPI; 1997-526451/48.
 DR N-PSDB; AAT89155.
 XX
 PT New isolated beta-1,3-glucanase enzyme - obtained from Oerskovia
 PT xanthineolytica, used particularly for the lysis of microbial cells
 PT for obtaining desirable products
 PS Example 2; Page 35-36; 64pp; English.

CC This polypeptide comprises a novel Oerskovia xanthineolytica (OX)
 CC enzyme that exhibits beta-1,3-glucanase (BG) activity. Its amino
 CC acid sequence was deduced from an isolated genomic DNA sequence
 CC (see AAT89155). Claimed DNA constructs that encode the novel BG (see
 CC also AAM29456 for corrected sequence), a mannose binding domain (see
 CC AAM29458) or a full-length enzyme, i.e. BG with mannose binding
 CC domain (see AAM29456), can be used to produce recombinant BG
 CC polypeptides, with or without a mannose binding domain, in fungal
 CC or bacterial host cells. BG polypeptides are used for the
 CC degradation or modification of beta-glucan containing material,
 CC especially for the gentle lysis of microbial cell walls, thereby
 CC enabling recovery of desirable intracellular products with a
 CC reduced amount of contaminants. They can also be used for the
 CC production of e.g. pigments, colourants, flavourants, yeast
 CC extracts, pharmaceuticals, food or feed compositions, and to
 CC prepare protoplasts for use in fusion, transformation and cloning
 CC studies.

SO Sequence 263 AA;

Query Match 15.2%; Score 340.5; DB 18; Length 263;
 Best Local Similarity 28.5%; Pred. No. 1.1e-23;
 Matches 103; Conservative 30; Mismatches 95; Indels 133; Gaps 13;

QY 26 IWMODEFDYFDG-----AKMOHEVTATGGNSEFOLYTODGANSFVRDGLFKFTLLAD 80
 DB 5 LMSDEPFGAAGSAPNPAVMNHETGAHGMNAELQNTTASRANSAL-DGQ----- 53
 QY 81 NINPQTGAPEFGTDFMYNGVLDMVAMYGACTNTDNGCYRTGAAGDIPPMASARVTFQKY 140
 DB 54 -----GNLVITA-----RREGDSY-----TSARMTTQGY 79
 QY 141 SFTHGRVYVAAKMPVGMVLPALWMLPEDWYTGMPRSGEIDIIETIGNRPFKNTGGEFL 200
 DB 80 QPOYGTIEARIQIPROGQIMPAFWMLGSGFPPTPSSGSDIDIMENVGFEPRHVG---- 134
 QY 201 GIOKMGSTMHGPGMDNRWLTSLPKHSDMANYGDNFHFPMWDSNGRLEFVNDENQA 260
 DB 135 -----TVH-GEYSGGS-GITGMTOHPQGSFADTFHFPAVDMKPELTFWVD--GQO 183
 QY 261 LLDVPEPLIDANPMWVDFEWEMGKFWLPQYENDNPMAGTNLAEFDQNFHFLNVAVGSTN 320
 DB 184 FHRVTRASVGANAM-----VFDDPEFLILNVAVGQW 215
 QY 321 GFIPDGCINRGDPALOKPMNSGNDWYNDAMRKFFDARGNMKWTVWDESDNNAQVDYIRV 380
 DB 216 PGYPDG-----TTQLP-----OQMKVDYVRV 236
 QY 381 Y 381
 DB 237 Y 237

RESULT 6

AAR89136
 ID AAR89136 standard; protein: 455 AA.
 AC AAR89136;
 XX
 DT 22-AUG-1996 (first entry)
 XX
 DE Bombyx mori mature LPS-binding protein.
 XX
 KW lipopolysaccharide binding protein; Bombyx mori; haemolymph; primer;
 KM Enterobacter cloacae; PCR; polymerase chain reaction; amplification;
 KW probe; LPS-complex; septic shock; injection; transgenic plant; vine;
 KM tobacco; tomato; potato; fungal infection; fungus.
 OS Bombyx mori.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 170
 FT /note= "putative N-glycosylation site"
 XX
 PN FR2721032-A1.
 XX
 PD 15-DEC-1995.
 XX
 PF 09-JUN-1994; 94FR-0007083.
 XX
 PR 09-JUN-1994; 94FR-0007083.
 XX
 PA (INSP) INST PASTEUR.
 XX
 PI Brey PT, Lee W;
 XX
 DR WPI; 1996-060094/07.
 XX
 PT New protein from Bombyx mori that binds bacterial
 PT lipopolysaccharide - esp. used to treat septic shock, also DNA
 PT encoding it, for producing transgenic plant(s) resistant to fungal
 PT attack
 XX
 PS Claim 4; Page 41-42; 53pp; French.
 XX
 CC This is the amino acid sequence of a mature lipopolysaccharide (LPS)
 CC binding protein from Bombyx mori. The protein was isolated from the
 CC haemolymph of fifth stage B.mori larvae after injection with
 CC Enterobacter cloacae strain 57-9. Partial amino acid sequence was used
 CC to generate PCR primers (AAT10280-1). These amplified a fragment of the
 CC gene used as a probe to obtain the full length gene (AAT10879) by
 CC screening a cDNA library derived from B.mori previously injected with
 CC heat-killed E.cloacae. The protein or fragments of it, can be used to
 CC treat diseases associated with LPS complexes e.g. septic shock, to remove
 CC LPS from products used for injection and to protect transgenic plants
 CC e.g. vines, tobacco, tomato or potato, against fungal infections.
 CC
 CC Sequence 455 AA:
 SQ
 Query Match 14.0%; Score 314; DB 17; Length 455;
 Best Local Similarity 27.3%; Pred. No. 6,8e-21;
 Matches 101; Conservative 58; Mismatches 143; Indels 68; Gaps 17;
 QY 27 VQDEDFYFDGAKWQ-HEVTATGGNGSEFQLYTQDGANSFVRDGLKIRPTLLADNINPQ 85
 DB 135 IFEBOFDSIDENVWQIEQYIPYHPEYFVSQYRNMLVSTADGNLHI-----NAKIQ 187
 QY 86 TGAP-FGTDPMVNGVLDVWAMGACNTNNGCYRGAAGDT-PPMASRVRKTFQKYSTP 143
 DB 188 OHMPEFLDSDISTYSGTLN--LFSGGTSS-AEACIKQASGADILPIVSGRI-TSISGAFPT 242
 QY 144 HGRVVVAAKMPYGDMLPAILMPLPDMVYGGWP-RSGEIDITETIGN-----RDFKMT 195
 DB 243 YGTVEIRAKLPGQDMLYPIELPLEPLKATGSNMYSGVYKIKACAKNAELYSQPNYSNT 302
 QY 196 GGEFLGIGKMGSTMHMGPGMD-DNRXYWLTSLPKHSDMNYGDNFTFWFDMSPNGIRFV 254

DB 303 -----VLXGPIIMDECRENLSKRRRDGTSWGDSPFTTYSQVMTPDFIALSV 350
 QY 255 DENQALIDVYPLIDANPMWVDFWEMWCKPMLPOYENDNPVAGTNLAPFDQNFETILNV 314
 DB 351 DGEEMARVEAPRDAL-----PAVCAHAPRHLLQAGSQMAPFDHFETITLGV 396
 QY 315 AVGTNGFIPDGCINRGDPAIQRKPMWGDYRNDAMR-----FPDARGNWKTWDEGDN 370
 DB 397 AAGGITEP-RDGSTSGG---VTRKPW-----RDSARASVHFWRHMSDWEPRWSQP--- 443
 QY 371 NAMQVDYIRV 380
 DB 444 -SLIVDEVKV 452
 RESULT 7
 AAR89137
 ID AAR89137 standard; Protein: 467 AA.
 AC AAR89137;
 XX
 DT 22-AUG-1996 (first entry)
 XX
 DE Bombyx mori full length LPS-binding protein.
 XX
 KW lipopolysaccharide binding protein; Bombyx mori; haemolymph; primer;
 KM Enterobacter cloacae; PCR; polymerase chain reaction; amplification;
 KW probe; LPS-complex; septic shock; injection; transgenic plant; vine;
 KM tobacco; tomato; potato; fungal infection; fungus.
 OS Bombyx mori.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..12
 FT /note= "signal peptide"
 FT Peptide 13..467
 FT /note= "mature protein"
 FT Modified-site 182
 FT /note= "putative N-glycosylation site"
 XX
 PN FR2721032-A1.
 XX
 PD 15-DEC-1995.
 XX
 PF 09-JUN-1994; 94FR-0007083.
 XX
 PR 09-JUN-1994; 94FR-0007083.
 XX
 PA (INSP) INST PASTEUR.
 XX
 PI Brey PT, Lee W;
 XX
 DR WPI; 1996-060094/07.
 DR N-PSDB; AAT10279.
 XX
 PT New protein from Bombyx mori that binds bacterial
 PT lipopolysaccharide - esp. used to treat septic shock, also DNA
 PT encoding it, for producing transgenic plant(s) resistant to fungal
 PT attack
 XX
 PS Claim 5; Page 42-43; 53pp; French.
 XX
 CC This is the amino acid sequence of the full length lipopolysaccharide
 CC (LPS) binding protein from Bombyx mori. The protein was isolated from
 CC the haemolymph of fifth stage B.mori larvae after injection with
 CC Enterobacter cloacae strain 57-9. Partial amino acid sequence was used
 CC to generate PCR primers (AAT10280-1). These amplified a fragment of the
 CC gene used as a probe to obtain the full length gene by screening a cDNA
 CC library derived from B.mori previously injected with heat-killed
 CC E.cloacae. The protein or fragments of it, can be used to treat diseases
 CC associated with LPS complexes e.g. septic shock, to remove LPS from
 CC products used for injection and to protect transgenic plants e.g. vines,
 CC tobacco, tomato or potato, against fungal infections.

XX	Sequence	467 AA:
XX	Query Match	14.0%; Score 314; DB 17; Length 467;
XX	Best Local Similarity	27.3%; Pred. No. 7e-21;
XX	Matches 101; Conservative	58; Mismatches 143; Indels 68; Gaps 17
QY	27 VMODEFDYFDGAKWO-HEVTATGGNSEFQLYTQDANSFVBDGKLFYPTLLADNINPQ	85
DB	147 IFEQFDSDIDENVMQIEQYIPFIHPEYFVSQYRNKLTFTSTADGNLHI-----NAKLQ	199
QY	86 TGAP-FGSDPFMYNGVLDMVWAMYGACTNTDNNCCYRTGAAGDI-PPAMSAHYTFQKYSPT	143
DB	200 QHMFETLDSITSGTLN--LPSGCTSS-AEACIKQASGADILPPIVSGRI-TSIGFAPT	254
QY	144 HGRVYVNAKMFVGDMLMPAIWMLPBDWYVGMP-RSGEIDLIETIGN-----RDFKNT	195
DB	255 YGTYEIRAKLLPGQDMLYELIIEPLFKKYGSSNVAAGYKICACAGNMLIYSGPNYSMT	314
QY	196 GGEFLGICOKMSYTMHWPGMD-DNRYMLTSLPKHSDMWNYGDNFHTFWFDWSPNGILREFV	254
DB	315 -----VLYGCPIMDECRENFLESTKRKRDGYSWGSDFHTYSVQWTPDFIALSV	362
QY	255 DDENALLDVPYPLIDANPFWYDFEWKGRPLPYQYENDNPAGGNTLAFDQNFHILNV	314
DB	363 DGEELARVEAPRDAL-----PACCAHAPRLLQAGSOMAFEDHFFYITLGV	408
QY	315 AVGGNGEIPPGCINRGSDPALOKPMSNGDWYNDAMRK---FPDARGMVKWTWDEGDN	370
DB	409 AAGGITEE-RGGSITSGG-VTKFW-----RDSARRKASVTFPKHMSDMPFRMSQP----	455
QY	371 NAMOVDTIRV 380	
DB	456 -SLIVDFYKV 464	
XX	RESULT 8	
XX	AAR11599	
XX	ID AAR11599 standard; Protein; 261 AA.	
XX	AAR11599;	
XX	25-MAR-2003 (updated)	
XX	18-JUN-1991 (first entry)	
XX	Beta-1,3-glucanase.	
XX	Alkalophilic; heat resistant.	
XX	Bacillus sp.	
XX	JP03053883-A.	
XX	07-MAR-1991.	
XX	20-JUL-1989; 89JP-0185928.	
XX	20-JUL-1989; 89JP-0185928.	
XX	(SHKJ) SHINGIJUTSU KAIHATSU KK.	
XX	WPI; 1991-113290/16.	
XX	N-PSDB; AA011293.	
XX	Heat resistant beta-1,3-glucanase gene DNA - derived from	
XX	alkali-compatible Bacillus sp.	
XX	Claim 7; Fig 2; 11pp; Japanese.	
XX	The new enzyme is derived from alkalophilic Bacillus Sp. The	
XX	enzyme acts on beta-1,3-glucans to form glucose and laminaribiose.	
XX	It is stable over a wide pH range with opt. activity at weakly	
XX	acidic pH and has good heat resistance.	

CC	XX	SO	Sequence	261 AA;	13.8%; Score 310; DB 12; Length 261;	26.9%; Pred. No. 7,8e-21;	Matches 100; Conservative 35; Mismatches 83; Indels 154; Gaps 17;
CC	XX	SO	Query Match				
CC	XX	SO	Best Local Similarity				
CC	XX	SO	Matches 100; Conservative 35; Mismatches 83; Indels 154; Gaps 17;				
CC	XX	SO	24 YHIWODEF--DYDGAQHHEV-TATGG-GNSEFOLYTODGANSFVDFDKLIFPTLLA 79				
CC	XX	SO	4 WSLVMSDEFNQNSLNPANWTAIEIGTSGSGMGNNELQYTSRPQNLQVSGGNLII--TAQR 61				
CC	XX	SO	80 DNINPQIGAPPGCTPMKNGVLVDWAMVYACNTINDNGCYRGAADIDIPMSARVRIQK 139				
CC	XX	SO	62 ES-----YGMNNT-----SARIKTGL 79				
CC	XX	SO	140 YSTHGRVYVAKKPVGDMLPAILMPLBEDWVYGGMPRSGEIDLIETIGNRDFKNTGGEF 199				
CC	XX	SO	80 QSFYRGKTEAIRKIPRSGGLPAPWMLGDEDSVSGMPFCGEIDIMERVNNPHV----- 134				
CC	XX	SO	200 LGIOKMGSTMTMGQWMDNRYWLTSLPKHSDMWNG-----DNFHTFWDMSPNGLR 251				
CC	XX	SO	135 -----GYVA---MDAN-----GHAE---YGRTSGLNDFSQYHTYSVEWEPNYIR 172				
CC	XX	SO	252 FFVVDENQALLDVEYPLIDANPMMWYDFEWENKPKPLPYENDNPAGGT-NLAPFQNFHF 310				
CC	XX	SO	173 WEFVGG-----VOYNEFYI-----ANGTGNTEEFQRPFFI 201				
CC	XX	SO	311 ILNVAVGSTNGFIPDGCINRGDPPALOKPWSNGDWYNDAMRKFPDARGNMKWTWDEGDN 370				
CC	XX	SO	202 ILNLAVGG-----NMPGSRNATIP-----FFAQ----- 224				
CC	XX	SO	371 NAMQVDYIRYVK 382				
CC	XX	SO	225 --MLVDYVRYQ 234				
CC	XX	SO	RESULT 9				
CC	XX	SO	AAB99272				
CC	XX	SO	AAB99272 standard; protein: 384 AA.				
CC	XX	SO	AAB99272;				
CC	XX	SO	11-SEP-2001 (first entry)				
CC	XX	SO	Bacillus circulans beta-1,3-glucanase.				
CC	XX	SO	beta-1,3-glucanase; enzyme; stockbreeding; bread manufacture; brewing.				
CC	XX	SO	Bacillus circulans.				
CC	XX	SO	JP2001120280-A.				
CC	XX	SO	08-MAY-2001.				
CC	XX	SO	01-NOV-1999; 99JP-0311073.				
CC	XX	SO	01-NOV-1999; 99JP-0311073.				
CC	XX	SO	(MEIJ) MEIJ SEIKA KAISHA LTD.				
CC	XX	SO	WPI; 2001-364762/38.				
CC	XX	SO	DR N-PSDB; AAH46359; AAH46360.				
CC	XX	SO	Beta-1,3-glucanase and beta-1,3-glucan-binding peptide and their genes				
CC	XX	SO	Claim 1; Page 8-9; 12pp; Japanese.				
CC	XX	SO	The present sequence is beta-1,3-glucanase from Bacillus circulans. The				
CC	XX	SO	protein can be used in various industrial fields such as stockbreeding,				
CC	XX	SO	cake and bread manufacture and brewing.				

[illegible]


```

Db      248 MFPEVPPYLLINAGGACMPAITRNADJ-----PFS----- 278
OY      361 KWTWDEGDNNAMQVDYIRYK 382
Db      279 -----MQVDYIRYK 288

RESULT 13
ABB60444
ID      ABB60444 standard; Protein: 410 AA.
XX
AC      ABB60444;
XX
DE      26-MAR-2002 (first entry)
XX
DE      Drosophila melanogaster polypeptide SEQ ID NO 8124.
XX
KM      Drosophila: developmental biology; cell signalling; insecticide;
XX      pharmaceutical.
XX
OS      Drosophila melanogaster.
XX
PN      MO200171042-A2.
XX
PD      27-SEP-2001.
XX
PF      23-MAR-2001; 2001WO-US09231.
XX
PR      23-MAR-2000; 2000US-191637P.
XX      11-JUL-2000; 2000US-0614150.
XX
PA      (PEKE ) PE CORP NY.
XX
PI      Venter JC, Adams M, Li PWD, Myers EW;
XX      N-PSDB; ABL04547.
XX
DR      WPI: 2001-656860/75.
XX
PT      New isolated nucleic acid detection reagent for detecting 1000 or more
XX      genes from Drosophila and for elucidating cell signalling and cell-cell
XX      interactions -
XX
PS      Disclosure: SEQ ID NO 8124; 21pp + Sequence Listing; English.
XX
CC      The invention relates to an isolated nucleic acid detection reagent
XX      capable of detecting 1000 or more genes from Drosophila. The invention is
XX      useful in developmental biology and in elucidating cell signalling and
XX      cell-cell interactions in higher eukaryotes for the development of
XX      insecticides, therapeutics and pharmaceutical drugs. The invention
XX      discloses genomic DNA sequences (AB101840-AB16175) and the encoded proteins
XX      (AB87737-AB872072).
XX      The sequence data for this patent did not form part of the printed
XX      specification, but was obtained in electronic format directly from WIPO
XX      at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ      Sequence 410 AA;

Query Match      11.1%; Score 249.5; DB 22; Length 410;
Best Local Similarity 24.5%; Pred. No. 6.9e-15;
Matches 91; Conservative 52; Mismatches 143; Indels 85; Gaps 14;

OY      23 QYHIWQDEFD--YFDGAKWQHEVTATGGSSEFOLYTODGA-NSFVRDGLFTKPTLLA 79
DB      111 OGELIPEDNFSEAKLKTWKHDIRQMYVEEELVAFDAANRCYKBEELHVPPIAT 170
OY      80 DNINPOTGAPFGTDEPNYNGVLDVWAMYGACTNTDNNGCYRTGAADI---PMAASRVR 135
DB      171 E-----VYDGSFK-----LGDCTAVESPQECNINHGIFYSIKPVSQAQIH 213
OY      136 TFOKYSFTHGRVYVNAHKMPVGDMLPAIMMLPDDWYGGGPRSGEIDITITGNRDPK-N 194

```

```

Db      214 TRNSESFKEGKIYVRAKLPKGDMLPEPYLMQPVSY-TYAEYHAKQLRIAYARGNANLRTK 272
OY      195 TGGEFLGIQKMGSTHMGPGMNDNRMYLTLSPKHSDDMYGDNFHFEMWDMSPNGLRFFY 254
Db      273 QGDDISGNHLIYGGGVVHHG-----NAVQFLDKISNSHGDGFHNTYIMORDKITLWY 327
OY      255 DDENQALLDVPYPLIDANPMWVDFWEMGKRWLPQYENDNPWAGTNLAFDQNFHFLNV 314
Db      328 DDE-----YVGYLDGLPF-----FNEKCFIIFGV 352
OY      315 AVGCTNGTIFPDGCTIRGGDPAL-----QKFWSGDWTNDAMKRFEDARGWKKWTFDDEGDN 370
Db      353 TVGGFLNF-----DDSLARDVYKRNRE--PRAALSFWOHRDAMAPTW---GRH 397
OY      371 NAMQVDYIRYK 381
Db      398 SAMVYDYIRYV 408

RESULT 14
ABB60451
ID      ABB60451 standard; Protein: 410 AA.
XX
AC      ABB60451;
XX
DE      26-MAR-2002 (first entry)
XX
DE      Drosophila melanogaster polypeptide SEQ ID NO 8145.
XX
KM      Drosophila: developmental biology; cell signalling; insecticide;
XX      pharmaceutical.
XX
OS      Drosophila melanogaster.
XX
PN      MO200171042-A2.
XX
PD      27-SEP-2001.
XX
PF      23-MAR-2001; 2001WO-US09231.
XX
PR      23-MAR-2000; 2000US-191637P.
XX      11-JUL-2000; 2000US-0614150.
XX
PA      (PEKE ) PE CORP NY.
XX
PI      Venter JC, Adams M, Li PWD, Myers EW;
XX      N-PSDB; ABL04554.
XX
DR      WPI: 2001-656860/75.
XX
PT      New isolated nucleic acid detection reagent for detecting 1000 or more
XX      genes from Drosophila and for elucidating cell signalling and cell-cell
XX      interactions -
XX
PS      Disclosure: SEQ ID NO 8145; 21pp + Sequence Listing; English.
XX
CC      The invention relates to an isolated nucleic acid detection reagent
XX      capable of detecting 1000 or more genes from Drosophila. The invention is
XX      useful in developmental biology and in elucidating cell signalling and
XX      cell-cell interactions in higher eukaryotes for the development of
XX      insecticides, therapeutics and pharmaceutical drugs. The invention
XX      discloses genomic DNA sequences (AB101840-AB16175) and the encoded proteins
XX      (AB87737-AB872072).
XX      The sequence data for this patent did not form part of the printed
XX      specification, but was obtained in electronic format directly from WIPO
XX      at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ      Sequence 410 AA;

Query Match      11.1%; Score 249.5; DB 22; Length 410;
Best Local Similarity 24.5%; Pred. No. 6.9e-15;
Matches 91; Conservative 52; Mismatches 143; Indels 85; Gaps 14;

```

```

QY 23 QYHIVODEFD--YFDGAKMOHEVYATGCGNSEFQLYTODGA-NSFYRDGKLFIPKPTLLA 79
DB 111 QGELIEEDNFSQAQINKTWKHDIRQMYHYEEELVAFDDAARACFYKEGELHIVPTIAT 170
QY 80 DNINPQTGAFPGTDFEMNGVLDVWAMTGACTNTDNNCCYRGAAGDI---PPMASARVR 135
DB 171 E-----VTDGSKF-----IGDRCTAVESPQECNIAHGIFYSIKPPVESAQIH 213
QY 136 TFOKYSFTHGRVYVYHAKMPYGDMLPAIMLPEDWYGGMPRSGEIDIIETIGRDKR-N 194
DB 214 TRNSFSFKRKIYVRAKLPKGDMLFPYLMQPV-S-TAETHYAKQINAYARGNANRTK 272
QY 195 TGGEFLGIQKMGSTMHGPGMDNRYWLTSLPKHSDDMYGDNEHTFWFMSPGRLREFV 254
DB 273 QGDIDSGNHLYGCGVYVHHG-----NAVQFLKDKISNSHYGDDEHNTMTMQRKITLMV 327
QY 255 DDENQALLDVPYPLIDANPWWVDPEWEGKFWLPQYENDNPWAGSTNLAPFDQNFHFLNV 314
DB 328 DDE-----VYGLYDGLPF-----FNEKCFIIFGV 352
QY 315 AVGTGTFIPDGCINRGDPAL---OKPWSNGDWYNDAMRKFPDARGNMKWTDDGDGN 370
DB 353 TYGGLNF-----DQSLAKDYKPYKRE--PRALSEFQHDAMAPTW--GRH 397
QY 371 NAMQVDYIRYV 381
DB 398 SAMVIDYIRYV 408

RESULT 15
AAR67915
ID AAR67915 standard; Protein; 673 AA.
AC AAR67915;
XX
XX
DT 25-MAR-2003 (updated)
DT 14-SEP-1995 (first entry)
XX
DE (1-3)-beta-D-glucan sensitive factor.
XX
XX (1-3)-beta-D-glucan sensitive factor; antifungal agent;
KM mycosis diagnosis.
XX
XX Limulus sp.
OS
XX
XX
FH Key Location/Qualifiers
FT Peptide 1..19
FT /label= sig_peptide
XX
XX W09501432-A1.
PN
XX
XX 12-JAN-1995.
PD
XX
XX 29-JUN-1994; 94MO-JP01057.
PF
XX
XX 29-JUN-1993; 93JP-0184403.
PR
XX
XX (SEGR ) SEIKAGAKU KOGYO CO LTD.
PA
XX
XX Iwanaga S, Muta T, Oda T, Seki N,
PI
XX
XX WPI: 1995-060996/08.
DR N-PSDB; AAO81334.
XX
XX
XX DNA encoding a polypeptide comprising a tetrapeptide motif at
PT least once - which may be used as an antibacterial and
PT antifungal.
XX
XX
XX Claim 9: Pages 23-32; 51pp; Japanese.
PS
XX
XX AAO81334 encodes AAR67915 a (1-3)-beta-D-glucan sensitive factor, it
CC has a high affinity for the (1-3)-beta-D-glucan found in fungal

```

```

CC cell walls. The protein is therefore useful for clinically
CC diagnosing mycosis, and as an antifungal agent for the removal
CC of fungi.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 673 AA;
SQ
Query Match 9.3%; Score 209; DB 16; Length 673;
Best Local Similarity 21.8%; Pred. No. 8.5e-11;
Matches 85; Conservative 49; Mismatches 106; Indels 150; Gaps 15;

QY 5 LVVL-CILFGRGFA---FTDQYHIVODEFDYFDGAKMOHEV--TATGCGNSEFQLYT 58
DB 2 LVLLCCVYLAHVVARICCSHEPKQQLVMSDEFTNGISSDWEPEMGNGLNGMNGNNLQYRR 61
QY 59 QDGNNSYVRDCKLEIKPFLADNINPQTGAFPGTDFEMNGVLDVWAMTGACTNTDNNCCY 118
DB 62 RE--NAQVEGSKLYI-----TAKREDYDGFK 85
QY 119 RTGAAGDIPPMASARVTFQKYSFTHGRVYVYHAKMPYGDMLPAIMLPEDWYGGMPRS 178
DB 86 YT-----SARLKTQFPKSKYKTEKKAIPSRGYVWYFWMGSDVNTVYRWPS 135
QY 179 GEIDIIETIGNRDEKNTGGEFLGIQKMGSTMHGPGMDNRYWLTSLP-----KHSDDMN 233
DB 136 GEIDFIE-----HRTNNE-----KVNGTTHW-----STPDGAHAHHNRSEN 172
QY 234 -YGDNEHTFWEFMSPGNLRFPVDENQALLDVPYPLIDANPWWVDPEWEGKFWLPQYEND 292
DB 173 TNGIDYHIYSVENNSSIVKWEVNGN-----QYFEV 202
QY 293 NPMAGSTNLAPFDQNFHFLNVAVGTGTFIPDGCINRGDPALOKPWSNGDWYNDAMRK 352
DB 203 KIOGCVNGKSAFRKRVVILMAIGN-----WPGFVADAE--- 239
QY 353 FPDARGNMKWTWDEGDGNNAMQVDYIRYV 382
DB 240 -FPAK-----MYIDYVRYQ 253

```

Search completed: September 16, 2003, 11:25:56
Job time: 120.038 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:23:52 ; Search time 40.6247 Seconds
(without alignments)
399,938 Million cell updates/sec

Title: US-09-596-101C-3

Perfect score: 2240
Sequence: 1 MKMTLVVLCILFGESEFAFTD.....DDEGNNAMQVDYIRVYKRN 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PCTRUS.COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	357	15.9	321	2	US-08-712-072C-3
2	351.5	15.7	262	1	US-08-392-828C-37
3	351.5	15.7	262	3	US-09-330-945-37
4	350	15.6	303	3	US-09-159-106-13
5	350	15.6	433	3	US-09-159-106-11
6	344.5	15.4	276	2	US-08-712-072C-4
7	340.5	15.2	263	3	US-09-159-106-2
8	307	13.7	306	2	US-08-824-707-2
9	286.5	12.8	285	2	US-08-712-072C-5
10	272.5	12.2	422	2	US-08-712-072C-2
11	208	9.3	634	1	US-08-392-828C-2
12	208	9.3	654	1	US-09-330-945-2
13	129.5	5.8	292	2	US-08-737-526-4
14	129.5	5.8	292	3	US-09-098-580-4
15	112.5	5.0	738	1	US-07-985-458-3
16	111.5	5.0	622	4	US-09-311-626B-4
17	110	4.9	243	3	US-09-286-690-10
18	107.5	4.8	545	3	US-09-269-731-8
19	106	4.7	238	3	US-09-286-690-7
20	101	4.5	465	4	US-09-658-772-2
21	100	4.5	237	1	US-08-103-998-4
22	99.5	4.4	1231	4	US-08-714-741-41
23	99	4.4	279	3	US-09-286-690-9
24	99	4.4	539	4	US-09-719-402A-2
25	98.5	4.4	239	1	US-08-103-998-2
26	98.5	4.4	308	4	US-09-463-862A-1
27	96.5	4.3	685	4	US-09-252-991A-32033

28	96	4.3	478	4	US-09-107-532A-4922	Sequence 4922, Ap
29	96	4.3	814	4	US-09-486-072-1	Sequence 1, Appl
30	95	4.2	1722	4	US-09-194-612A-1	Sequence 1, Appl
31	94.5	4.2	276	4	US-09-719-402A-6	Sequence 6, Appl
32	94.5	4.2	1278	4	US-09-604-957-3	Sequence 3, Appl
33	93.5	4.2	1052	4	US-09-360-237-1	Sequence 1, Appl
34	93	4.2	829	4	US-09-252-991A-27150	Sequence 27150, A
35	92.5	4.1	312	3	US-09-104-308-1	Sequence 21, Appl
36	92.5	4.1	371	3	US-09-321-981-1	Sequence 1, Appl
37	92.5	4.1	371	3	US-09-739-861A-1	Sequence 1, Appl
38	92.5	4.1	371	4	US-09-795-583-1	Sequence 1, Appl
39	92.5	4.1	386	4	US-09-321-981-5	Sequence 5, Appl
40	92.5	4.1	386	3	US-09-739-861A-5	Sequence 5, Appl
41	92.5	4.1	386	4	US-09-795-583-5	Sequence 5, Appl
42	92.5	4.0	242	3	US-09-286-690-11	Sequence 11, Appl
43	90.5	4.0	429	1	US-08-745-977-4	Sequence 4, Appl
44	90.5	4.0	429	3	US-09-040-699A-4	Sequence 4, Appl
45	90.5	4.0	429	3	US-09-040-699A-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-712-072C-3
Sequence 3, Application US/08712072C
Patent No. 5925541
GENERAL INFORMATION:
APPLICANT: Jack Goldstein, Alex Zhu and Lin Leng
TITLE OF INVENTION: ENDO-BETA-GALACTOSIDASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amster, Rothstein & Ebenstein
STREET: 90 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE DISKETTE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/712,072C
FILING DATE: 11-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Elizabeth A.
REGISTRATION NUMBER: 39,911
REFERENCE/DOCKET NUMBER: 63475/97
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: el3b, Bacillus circulans
US-08-712-072C-3
Query Match 15.9%, Score 357, DB 2, Length 321;

Best Local Similarity 26.2%; Pred. No. 1.2e-27;
Matches 100; Conservative 46; Mismatches 86; Indels 150; Gaps 13;

```

OY      25  HVMWDEF-----DYPGAKQKHVETATGGNSEFOLYTOGANSFYADKLFK  74
          ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      64  NLIWODDERETTLDTSKMYETGYLLNDRPALMGWMAELQHTNSTONYVYDGLKIK  123
          ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY      75  PTLADNINPQTGAPEFGTDEMYNGVLVDVWAMYGACATNTDNNCGYRTGAAGDIPANSARV  134
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      124  --AMNDSASPOPR-----NRAAQS-----SGKI  145
          ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY      135  RTFOKYSTHGRVYVYVNAKMPGVGDWMLPAILMILPEWVYVGGWPRSGEIDILITIGNRDFKN  194
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      146  NTKDKLSIKYGVYDVRPRAKLPRTGDGVWALMMLPKDYSYGTWAAASGEIDVWEARGRLPGSV  205
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      195  YTGGEFLGIQKMGSTTHWVGCGMDNRKRWLSLRKHSDDN-----YGDNRHTTWEWSPN  248
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      206  SG-----THFGGQMPVNO-----SSGGDYHFRREGQTEANDYHVSUYVEED  247
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      249  GLRFEVUDE-----NOALLDVPYPLIDANPMWVDEWEMGCKPWLPOYENDNPAWGSTNLA  302
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      248  NIKWVYDDGKEFYKVINQ-----QWTSIAPANNPN-----A  277
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      303  PPDONHFLILNVAVGST--NGFIPOGCIIRGGDPALQKPMWSNGMDWYNDAWRKEFFDARGMW  360
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      278  PFDEFFYLLIMNLAVGNGNEDGGRTP---NASDIPA-----308
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      361  KMTWDEGDGNMAMQVDYIRYVK  382
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      309  -----TMQVDYVAVYK  319
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 2
 US-08-392-828C-37
 Sequence 37, Application US/08392828C
 Patent No. 5793962
 GENERAL INFORMATION:
 APPLICANT: IWANAGA, SADAKI
 APPLICANT: MUTA, TATSUSHI
 APPLICANT: SEKI, NORIKI
 APPLICANT: ODA, TOSHIO
 TITLE OF INVENTION: NOVEL POLYPEPTIDE AND DNA ENCODING
 TITLE OF INVENTION: THEREOF
 NUMBER OF SEQUENCES: 39
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
 ADDRESSEE: THIBAUTJT
 STREET: 53 STATE STREET
 CITY: BOSTON
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/392,828C
 FILING DATE: 28-FEB-1995
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: CAMPBELL, PAULA A
 REGISTRATION NUMBER: 32,503
 REFERENCE/DOCKET NUMBER: FJN-033
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 248-7000
 TELEFAX: (617) 248-7100
 INFORMATION FOR SEQ. ID NO: 37:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 262 amino acids
 TYPE: amino acid
 STRANDEDNESS: single

```

? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? FEATURE:
? NAME/KEY: Protein
? LOCATION: 1..262
? OTHER INFORMATION:
?
? OS=08-392-828C-37 /note="BGI AI SEQUENCE (FIGURE 2)"

```

Query Match	15.7%	Score 351.5	DB 1	Length 262
Best Local Similarity	26.2%	Pred. No. 3.2e-27		
Matches 100; Conservative	44;	Mismatches 89;	Indels 149;	Gaps 14

```

QY      25 HIWVODEED--YFQGAQKHQEV-----TATGGNSEFOLYUDDGANSFVROCKLEFK 74
Db      4  NLIMODEFNGTLLDTSKNRYETGYLLNDPRATWGMAGNAELOHTYNTSTONRYVODCKLEIK 63
QY      75 PTLIADINIPOTGAPFGTDFMYNGVLVDYMAAYGACTNTDNNGCYRTGAAGDIIPAMASARV 134
Db      64  A-----MNDKSPFPDP-----NNYAAQYS-----SCKI 86
QY      135 RTEQKYSFETGCRVVYVNAKMPYGVDLMLPALIMLPEDDWYGGMPRSGEIDIETIGNRDFKN 194
Db      87  NTKKRLSLTKGRVDFRAKRLPTGGDGVWALMMLPRDSYVGTWAAAGEIIVMEARGRLPESV 146
QY      195 TGGEFLDIQKKGSMHMGPGMDRMYMLTSLPKSDMM-----YGDNFTHEFMDSPN 248
Db      147  SG-----YIHFQGGMPVNO-----SSGGDGHPEPGQTFADNYHVSYYVVEED 188
QY      249 GLRFEVDE-----NOALLDVPYPLIDANPMWVDFEWEGKRPWLPQYENDPMWAGGTMLA 302
Db      189  NIKRYVQSKFFPKYTNO-----QWYTAAPNNNN-----A 218
QY      303 PFDONFHFILNAVAGT--NGEIPDGCINRGDPALOKPWSNGDWYNDAMRKFEDARGNW 360
Db      219  PFDEPFILIMLAVAGNFDGGRTP-----NASDIPA-----249
QY      361 KMTWDDGSDNNAMAVDYIRYK 382
Db      250  -----TQVDVYRYK 260

```

RESULT 3
US-09-330-945-37
Sequence 37, Application US/09330945
Patent No. 6077946
GENERAL INFORMATION:
APPLICANT: IWANAGA, SADAAKI
APPLICANT: MUTA, TATSUSHI
APPLICANT: SEKI, NORIAKI
APPLICANT: ODA, TOSHIO
TITLE OF INVENTION: DNA ENCODING HORSESHOE CRAB
TITLE OF INVENTION: AMBIOCYTE LYSATE FACTOR G SUBUNIT A
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
ADDRESSEE: THIBEAULT, LLP
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/330,945
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/119,995
FILING DATE:

```

; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER, EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: FJN-032DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 262 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..262
; OTHER INFORMATION: /note="BGI A1 SEQUENCE (FIGURE 2)"
US-09-330-945-37

Query Match 15.7%; Score 351.5; DB 3; Length 262;
Best Local Similarity 26.2%; Pred. No. 3.2e-27;
Matches 100; Conservative 44; Mismatches 89; Indels 149; Gaps 14;

QY 25 HIWQDEFD--YFDGAKQHEV-----TATGGNSEFQLYTQDGANSEYRQDKLFK 74
DB 4 NLHQDEFNGTTLDTSKMNETGYLLNDPATWGMGNELQHYNSTQNYVYQDGLKIK 63
QY 75 PTLADINPOTGAPFGTDFMYNGVLVWAMYGACTNDNNGCYRTGAAGDIPMSARV 134
DB 64 A-----MDSKSPQDP-----NRAYYS-----SKTI 86
QY 135 RTQOKYSFTIGRYYVYHAKMPYQDLMPALIMLPEDWYVGMPRSGEIDIIETIGNRDEKN 194
DB 87 NTDKRLSLKYGKDFRAKLPFGDGVMPALMMLPKDSYGTWASGEIDVMAERGLPGSV 146
QY 195 TGGEFLDIOKMGSTMHGPGMDNRYWLTSLPKHSDMN-----YGNFTFWFWSPN 248
DB 147 SG-----TIHGGQPVNQ-----SSGQDHPEEGOTFANDYHYVSWEED 188
QY 249 GLRFEVDE-----NQALLDVPRYPLIDANPMWVDFEWGKRPMLPYENDPMWAGTNLA 302
DB 189 NIKMYVDGKFFKYKTNQ-----QWYSTAPNPN-----A 218
QY 303 PFDONHFIILNAVAGT--NGFIDGGINRGDPALOKPMSNGDMYNDAMKFFDARGNW 360
DB 219 PFDEPFLLMNLAVGNGFDGGRF-----NASDIPA----- 249
QY 361 KMTWDEGDNNMAQVDIIRYK 382
DB 250 -----TWQVDVRYK 260

RESULT 4
US-09-159-106-13
; Sequence 13, Application US/09159106
; Patent No. 6284509
; GENERAL INFORMATION:
; APPLICANT: Ferrer, Pau
; APPLICANT: Diers, Ivan
; APPLICANT: Halkier, Torben
; APPLICANT: Hedegaard, Lisbeth
; TITLE OF INVENTION: An Enzyme with -1,3-Glucanase
; FILE REFERENCE: 4693.204-US
; CURRENT APPLICATION NUMBER: US/09/159,106
; EARLIER FILING DATE: 1998-09-23
; EARLIER APPLICATION NUMBER: 0427/96
; EARLIER FILING DATE: 1996-12-04
; EARLIER APPLICATION NUMBER: 0885/96
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: PCT/DK97/00160
; EARLIER FILING DATE: 1997-04-14

```

```

; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Oerskovia xanthineolytica
US-09-159-106-13

Query Match 15.6%; Score 350; DB 3; Length 303;
Best Local Similarity 28.5%; Pred. No. 5.6e-27;
Matches 103; Conservative 30; Mismatches 96; Indels 132; Gaps 12;

QY 26 IYWQDEFYFDG-----AKQHEVYATGGNSEFQLYTQDANSFVRDGLFKPILAD 80
DB 57. ILMSEDFQAGASAPNPAPVWNETGAGHGMNLEQVYTSRANSAL-DSQ----- 105
QY 81 NINPOTGAPFGTDFMYNGVLVWAMYGACTNDNNGCYRTGAAGDIPMSARVTFQKI 140
DB 106 -----GNLVITA-----RREGDSY-----TSARMTQGRK 131
QY 141 SFTHGRVYVYHAKMPYQDLMPALIMLPEDWYVGMPRSGEIDIIETIGNRDEKNTEGEFL 200
DB 132 QPQYRIARLQIDPRGQGWPAFPMWLGSEFPPTPMSGEIDIMENVGEPRHYHG----- 187
QY 201 GIQKMGSTMHGPGMDNRYWLTSLPKHSDMNNGDNFTFWFWSPNGLRFFVDENQA 260
DB 188 -----TVH-GPGYSGGS-GITGMVQHPOGWSFADTFHTFVDMKPGETTFVFD--GQQ 236
QY 261 LIDVPRYPLIDANPMWVDFEWGKRPMLPYENDNPMAGTNINLAPDONHFIILNAVAGTN 320
DB 237 FHRVTRASYGANAW-----VFDDPFLLILNAVAGGW 268
QY 321 GFIPDGINRGDPAIOLKPFWSNGDMYNDAMKFFDARGNMKWTWDEGDNNMAQVDIIRY 380
DB 269 PGYPDG-----TTQLP-----QQMKVDYVRK 289
QY 381 Y 381
DB 290 Y 290

RESULT 5
US-09-159-106-11
; Sequence 11, Application US/09159106
; Patent No. 6284509
; GENERAL INFORMATION:
; APPLICANT: Ferrer, Pau
; APPLICANT: Diers, Ivan
; APPLICANT: Halkier, Torben
; APPLICANT: Hedegaard, Lisbeth
; TITLE OF INVENTION: An Enzyme with -1,3-Glucanase
; FILE REFERENCE: 4693.204-US
; CURRENT APPLICATION NUMBER: US/09/159,106
; EARLIER FILING DATE: 1998-09-23
; EARLIER APPLICATION NUMBER: 0427/96
; EARLIER FILING DATE: 1996-12-04
; EARLIER APPLICATION NUMBER: 0885/96
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: PCT/DK97/00160
; EARLIER FILING DATE: 1997-04-14
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Oerskovia xanthineolytica
US-09-159-106-11

Query Match 15.6%; Score 350; DB 3; Length 435;
Best Local Similarity 28.5%; Pred. No. 9.4e-27;
Matches 103; Conservative 30; Mismatches 96; Indels 132; Gaps 12;

```


Query Match	13.7%;	Score 307;	DB 2;	Length 306;
Best Local Similarity	26.2%;	Pred. No. 1.2e-22;		
Matches	95;	Conservative	36;	Mismatches 97;
			Indels	134;
			Gaps	13

```

; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO

```

ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: Flavobacterium keratolyticus
US-08-712-072C-5

Query Match 12.8%; Score 286.5; DB 2; Length 285;
Best Local Similarity 27.0%; Pred. No. 1.2e-20;
Matches 103; Conservative 32; Mismatches 100; Indels 147; Gaps 15;

QY 10 LLEGGFAFTMDQYHIWODEFDY---FDGAKMOHEVTATGGN---SEFOLY-TQDGA 62
DB 8 LLNATVATTD---YELIMSDEFNSSGGFDSFKWSYADRGTVAMNKYMTSLPAVSODGS 64
QY 63 NSFYRDGKLFKPTLLADNINPQTGAPFOTDMYNGVLDYWMYAGACTNTDNNCCTRTGA 122
DB 65 NLVNL-----MDNAV----- 75
QY 123 AGDIPRMSARVTFPKYSFTHGRVYVHAKMPGMLPAILMPLPD--WYGGMPRSGEI 181
DB 76 AGDPYAHAGYKSMGKFSMTYKVEYRAKFTQSGSWPALIMMPEPATAYGWPSCGEI 135
QY 182 DIETIGNRDPEKNTGEEFLGIQKMGSTMHMGPGMDNRYWLSLPKHSDDMYGDNFHTF 241
DB 136 DSMEHVNNESV-----MYHTINGSVTNANG---GSTASKSATYNTTD-YNLY 179
QY 242 WFDWSPNLRFPVVDENQALLDVPYPLIDANPMWVDFWEMGKFWLPQYENDNPMAGCTYL 301
DB 180 TWIMSPNDIRYV---NNSL-----QYTYARVSGGCTQO 210
QY 302 APFDQNFHFLNVAVG-GTNGFLPDGGINRGDPALQKPMNSGNDWYNDAMRKFFDARGNW 360
DB 211 WPFDPFYLILNQAGAGMPGAIITNADL-----PFS----- 241
QY 361 KWTWDEGDNNMAOVYIRYK 382
DB 242 -----MQVDYVRVYK 251

RESULT 10

US-08-712-072C-2
; Sequence 2, Application US/08712072C
; Patent No. 5925541
; GENERAL INFORMATION:
; APPLICANT: Jack Goldstein, Alex Zhu and Lin Leng
; TITLE OF INVENTION: ENDO-BETA-GALACTOSIDASE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amster, Rothstein & Ebenstein
; STREET: 90 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,072C
; FILING DATE: 11-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Elizabeth A.
; REGISTRATION NUMBER: 39,911
; REFERENCE/DOCKET NUMBER: 63475/97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-5995
; TELEFAX: (212) 286-0854 or 286-0082

TELEX: TWX 710-581-4766
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 422 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; US-08-712-072C-2

Query Match 12.2%; Score 272.5; DB 2; Length 422;
Best Local Similarity 27.2%; Pred. No. 5.3e-19;
Matches 105; Conservative 29; Mismatches 97; Indels 155; Gaps 17;

QY 10 LLEGGFAFTMDQYHIWODEFDY---FDGAKMOHEVTATGGNSEFOLYTQDANSFV 66
DB 45 LLNATVATTD---YELIMSDEFNSSGGFDSFKWSYADRGTVAMNK--YMTSODGSNLVL 99
QY 67 ROGKLFKPTLLADNINPQTGAPFGTDFMYNGVLDYWMYAGACTNTDNNCCTRTGAADI 126
DB 100 R-----MDNAV-----AGD- 108
QY 127 PRMSARVTFPKYSFTHGRVYVHAK-----MPV-----GDMLPAILMPLPD--WYGGMPR 177
DB 109 PVAHYHAGYKSMGKFSMTYKVEYRAKFTGVSLEPAVQGRGSPALIMMPEPATAYGWP 168
QY 178 SCEIDIEITIGNRDEKNTGEEFLGIQKMGSTMHMGPGMDNRYWLSLPKHSDDMYGDN 237
DB 169 CEGIDSMEHVNNESV-----MYHTINGSVTNANG---GSTASKSATYNTTD- 212
QY 238 FHTWFDWSPNLRFPVVDENQALLDVPYPLIDANPMWVDFWEMGKFWLPQYENDNPMAG 297
DB 213 YNLYTWIMSPNDIRYV---NNSL-----QYTYARVSGG 243
QY 298 GTNLAPFDQNFHFLNVAVG-GTNGFLPDGGINRGDPALQKPMNSGNDWYNDAMRKFFDA 356
DB 244 GTQQPFYDVPFYLILNQAGAGMPGAIITNADL-----PFS----- 278
QY 357 RGNMKWTWDEGDNNMAOVYIRYK 382
DB 279 -----MQVDYVRVYK 288

RESULT 11

US-08-392-828C-2
; Sequence 2, Application US/08392828C
; Patent No. 5795962
; GENERAL INFORMATION:
; APPLICANT: IWANAGA, SADAOKI
; APPLICANT: MOTA, TATSUSHI
; APPLICANT: SEKI, NORIAKI
; APPLICANT: ODA, TOSHIO
; TITLE OF INVENTION: NOVEL POLYPEPTIDE AND DNA ENCODING
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

Query Match	9.38;	Score 208;	DB.1;	Length 654;
Best Local Similarity	21.58;	Pred. No. 2.9e-12;		
Matches 79;	Conservative 45;	Mismatches 98;	Indels 146;	Gaps 13

RESULT 12
US-09-330-945-2

```

1 GENERAL INFORMATION:
2 APPLICANT: IWANAGA, SADAOKI
3 APPLICANT: MUTA, TATSUSHI
4 APPLICANT: SEKI, NORIAKI
5 APPLICANT: ODA, TOSHIO
6 TITLE OF INVENTION: DNA ENCODING HORSESHOE CRAB
7 TITLE OF INVENTION: AMBIOCYTE LYSATE FACTOR G SUBUNIT A
8 NUMBER OF SEQUENCES: 39
9 CORRESPONDENCE ADDRESS:
10 ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
11 ADDRESSEE: THIBEAULT, LLP
12 STREET: 125 HIGH STREET
13 CITY: BOSTON
14 STATE: MA
15 COUNTRY: USA
16 ZIP: 02110
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: floppy disk

```

Query Match	9.3%;	Score 208;	DB 3;	Length 654;
Best Local Similarity	-21.5%;	Pred. NO. 2.9e-12;		
Matches 79;	Conservative 45;	Mismatches 98;	Indels 146;	Gaps 13

RESULT 13
US-08-737-526-4

APPLICANT: Kotof, Iene Venke
 APPLICANT: Andersen, Iene No. 5871966boc
 APPLICANT: Kauppinen, Markus Sakari
 APPLICANT: Christgau, Stepan
 TITLE OF INVENTION: An Enzyme With Endo-1,3(4)-B-Glucanase
 TITLE OF INVENTION: Activity
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 5871966o No. 5871966disk of No. 5871966th America, Inc
 STREET: 405 Lexington Avenue

```

CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Versio
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/737,526
FILING DATE: 08-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Valeta, Gregg A
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 4174.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-737-526-4

```

Query Match	5.8%	Score	129.5	DB	2	Length	292
Best Local	Similarly	22.0%	Pred.	NO. 7e-05			
Matches	85	Conservative	40	Mismatches	110	Indels	151
						Gaps	22

QY	24	YHIVQDEDEYEDGCA-----KQKHETATANGSEFOLYODCAFSFVRNGKLFKRTLL	78
Db	28	FNLVWTDYFAGGAGGGSPPQNNM--NITTGMLNNADEYTSSTSIANYQLSGS-----TL-	80
QY	79	ADNINPQTGAPEGTDFMYNGVLDVAMVYAGCTNTDNNGCYRTGAAGDIPRMSARVTEQ	138
Db	81	-QLVPMRDSKSGT-----STFGWT-----SGRLS--	105
QY	139	KYSEF--HGRRV-VIAKMPVG-----DMLPAMLMPEDWYV--GMPRSEIDIIET	166
Db	106	KYTFPPAAGKYATLEAAIRFGSMAQANROGIMPAPFMLDSDSLRPGSGSWPNCGEIDIMET	165
QY	187	IGNRPFKNTGGEEL-----GIQKNGSPMHNP--GMDNRMYLTLSPKSDDMNYG	235
Db	166	V---DGAATGHCGLCHDYVPGGICNEGNET--GGPVAIANYNDHNAFVETIDRTPSSV--	218
QY	236	DNFHTFMDWSPENGLRFVYDDENOALLDVPYPLIDANPMWVDEWEMGKPYLPQYENDNPV	295
Db	219	---QSEITLWSLDGLITFYQITGS-----RIGNGVW	246
QY	296	AGGTINLAFPDQNFHTILNVAVGCTNGEIPDGCINRGDPALQKFWNSGDMYNDAMRKFFD	355
Db	247	--NNIA--HSPLEFLILNVAVGG-----NMFGNP-----NSATL---	275
QY	356	ARGNNKWTWDEGCDNNANQVDIRYV	381
Db	276	-----DGYGSMMEVGIVAOI	290

```

RESULT 14
US-09-098-580-4
; Sequence 4, Application US/09098580
; Patent No. 6140096
;
; GENERAL INFORMATION:
;
; APPLICANT: Kofod, Lene Venke
;
; APPLICANT: Andersen, Lene No. 6140096boc
;
; APPLICANT: Kauppinen, Markus Sakari
;
; APPLICANT: Christgau, Stephan
;
; TITLE OF INVENTION: An Enzyme With Endo-1,3(4)-B-Glucanase
;
; TITLE OF INVENTION: Activity

```

NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 61400960 NO. 6140096disk of NO. 6140096ctt America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/0998,580
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/737,526
FILING DATE: 08-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Valeta, Gregg A
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 4174,204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-098-580-4

Query Match	5.8%	Score 129.5	DB 3	Length 292
Best Local Similarity	22.0%	Pred. No. 7e-05		
Matches 85	Conservative 40	Mismatches 110	Indels 151	Gaps 22

```

OY      24 YHIVQDEDFEDGDA-----KQWHEVUTATGGNSFOLYTOGANSFVRGKLFIRKPTLL 78
Db      28 FNLVMTDIFPAONGSGSPNQNMM-NIITGLNLNNABEITYSSSRANQVLSGGS-----TL- 80
OY      79 ADNINPOTGAPFGTDEMYNGVLVWAMYGACNTNDNNGCYRTGAAGDIPPMASARVYRTEQ 138
Db      81 -QLVPMRDSKGT-----STFGGWT-----SGRLS----- 1050
OY      139 KYSTF--HGRVY-VHAKMPVG-----DWLPMATMLPEWDVY--GMPRSEIDIIET 186
Db      106 KYTFPAAGKYTRLEAAIRFGSMAQANKOGIMPAPFMMLODSLRQPGSGMPNCEIDIMET 165
OY      187 IGNRFRKNTGGBEFL-----GIQKMGSTMHNGP-----GMDDNRVYWLTSLPKHSDDMYG 235
Db      166 V---DGOATGHTGLTCLDYYPGSICWEGNKI--GGPVAILANVNDHNAKVELIDRTPSSH--- 218
OY      236 DNFHTFWMDSPBNGILRFVDDENQALLDVPYPLIDANPMWYDWEWEGKFWLPQYENDNPM 295
Db      219 ---QSEITLWLSLDGITIYQITGS-----RIGNQVW 246
OY      296 AGGTNLAPFDONFHFILVANVGCTGCFIPDCCINRGGDALOKPWSNGMDYNDAMKRFED 355
Db      247 ---NNIA--HSLPFLILNVAVG-----NMPGNP-----NSATL----- 275
OY      356 ARGNNKWTWDEGDNNAQOVDYIRY 381
Db      276 -----DGTGSMEEVGIVAOY 290

```

RESULT 15
US-07-985-458-3
; Sequence 3, Application US/07985458
; Patent No. 5344777

GENERAL INFORMATION:
 APPLICANT: Tamaki, Toshimi;
 APPLICANT: Takemura, Hiroshi;
 APPLICANT: Takemura, Kenji;
 APPLICANT: Fukaya, Masahiro;
 APPLICANT: Okumura, Hajime and
 APPLICANT: Kawamura, Yoshiya
 TITLE OF INVENTION: Structural Gene of Membrane-Bound
 TITLE OF INVENTION: Alcohol Dehydrogenase Complex, Plasmid
 TITLE OF INVENTION: Containing The Same And Transformed Acetic Acid
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Frishauf, Holtz, Goodman & Woodward, P.C.
 STREET: 600 Third Avenue
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10016-2088
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 Inch, 0.72 mb
 COMPUTER: IBM PC compatible (NEC PC-9801 ES)
 OPERATING SYSTEM: MS DOS
 SOFTWARE: ASCII Form
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/985,458
 FILING DATE: 19921203
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/658,221
 FILING DATE: 20-FEB-1991
 APPLICATION NUMBER: 73440/1990
 FILING DATE: 26-MAR-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Goodman, Herbert
 REGISTRATION NUMBER: 17081
 REFERENCE/DOCKET NUMBER: 910134/HG
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212)972-1400
 TELEFAX: (212)370-1622
 TELEX: 236268
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 738 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: NATURE PEPTIDE
 LOCATION: 36 to 738
 IDENTIFICATION METHOD: N-terminal sequences of the
 IDENTIFICATION METHOD: purified protein having a molecular weight of about
 IDENTIFICATION METHOD: 72,000
 ORIGINAL SOURCE:
 ORGANISM: Acetobacter alcoacetiogenes
 STRAIN: MH-24
 PUBLICATION INFORMATION:
 AUTHORS: Tamaki, Toshimi;
 AUTHORS: Takemura, Hiroshi;
 AUTHORS: Takemura, Kenji;
 AUTHORS: Okumura, Hajime;
 AUTHORS: Kawamura, Yoshiya;
 AUTHORS: Nishiyama, Makoto;
 AUTHORS: Horinouchi, Sueharu and
 AUTHORS: Beppu, Teruhiko
 TITLE: Cloning and Sequencing of the Gene Cluster
 TITLE: Encoding Two Subunits of Membrane-Bound
 TITLE: Alcohol Dehydrogenase from Acetobacter
 TITLE: polyoxigenes
 JOURNAL: Biochimica et Biophysica Acta.
 VOLUME: 1088
 PAGES: 292-300

DATE: 1991
 US-07-985-458-3
 Query Match 5.0%; Score 112.5; DB 1; Length 738;
 Best Local Similarity 20.4%; Pred. No. 0.013;
 Matches 78; Conservative 37; Mismatches 113; Indels 155; Gaps 21;
 QY 48 GGNSEFQLYTQDGSFVR-----DGKLFKPTLLADNINPQTGAPGPD-FMYN 97
 Db 212 GNGGSEF-----GARGFVSFADATGKVDWRFETVP-----NPKNEPDAASDVLNN 258
 QY 98 GVLVAMNYGACTYTDNNGC-----YRTGAGD----- 125
 Db 259 KAVQWSPGTGAWTRQGGGTVWDSIYDPAVDLVYLGNGSPMNYKYSRGSGONFLG 318
 QY 126 ---IPRMSARVTFPO-----KYSFTHGR-----VVYHAK----- 152
 Db 319 SIYALKRPETGEYVYHFPETPDQWDFTSQOIMPLDLPINGETRHVIVHAKNGEFTIID 378
 QY 153 MPVGDWLMPIAWMLPEDWVYGCW-----PRSGEIDIIETIGNRD--EKNTGGEFLGID-K 204
 Db 379 AKTGEFT-----SGKNVYVYVNMASGLDPKGTG-----RPIYNPDALYTLTGKEWYGIKPD 427
 QY 205 MG---STMHMGPCMDNRYWLTSLP-----KHSDDMNYGDNFHTFED 244
 Db 428 LGGHNFAPAMAFSP-----KTGLVYIPAOQVPELYTNOVGFTPHDSDNNLGLDMKKVGIP 482
 QY 245 WSPNGLRFFVDDENQALLDVPYPLIDANPMWVDFWE--WGKPMLPQYENDNPMVAGTNL 301
 Db 483 DSPBAKQAFVKDLK-----GWIYVMDPQKQAEAM--RVDHKGPMNNGILA 525
 QY 302 APEDQNFHTLN---VAVGTNG 321
 Db 526 TGGDLPLQGLANGEFHAYDATING 548

Search completed: September 16, 2003, 11:30:26
 Job time : 42.6247 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:15:27 ; Search time 27.0831 Seconds
(without alignments)
666,771 Million cell updates/sec

Title: US-09-596-101C-3

Perfect score: 2240
Sequence: 1 MRWTLVYVCLLFESEGFATD.....DDESDNNAMQVDYIRYTKRN 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	351.5	15.7	682	1 E13B_BACCI	P23903 bacillus ci
2	344.5	15.4	286	1 GUB_RHOMR	P45798 rhodothermu
3	175.5	7.8	465	1 EXSH_RHME	O33680 rhizobium m
4	167	7.5	465	1 EGIC_RHME	Q92342 rhizobium m
5	112.5	5.0	738	1 DHER_ACBO	P28036 acetobacter
6	112.5	5.0	954	1 XYNB_ROMFL	P29126 rumiococcu
7	110	4.9	243	1 GUB_BACLI	P27051 bacillus li
8	109	4.9	781	1 NANB_VIECH	P37060 vibrio chol
9	108	4.8	720	1 KRE6_YEAST	P32486 saccharomyc
10	106	4.7	739	1 GUB_PABPO	P45787 paenibacilli
11	106	4.7	739	1 DHER_ACCEU	Q44002 acetobacter
12	106	4.7	782	1 OSTA_HAEIN	P48464 haemophilus
13	105	4.7	802	1 XYNB_ROMFL	O53317 rumiococcu
14	104.5	4.7	239	1 GUB_BACAM	P07980 bacillus am
15	103.5	4.6	740	1 KRE6_CANAL	P87023 candida alb
16	102.5	4.6	334	1 GUB_CLOTH	P29716 clostridium
17	102	4.6	1028	1 FDXG_HAEIN	P46448 haemophilus
18	99	4.4	737	1 SKNI_CLOTH	P87024 candida alb
19	99	4.4	1087	1 XYNB_CLOTH	P38535 clostridium
20	98	4.4	1196	1 AMVB_PABPO	P21543 paenibacilli
21	97.5	4.4	1157	1 XYNB_THESA	P36917 thermoaeroc
22	96.5	4.3	441	1 YI07_METUA	O60306 methanococc
23	96.5	4.3	623	1 EXAA_PSEAE	O92417 pseudomonas
24	96	4.3	728	1 GLGB_PCOLI	P07762 escherichia
25	96	4.3	771	1 SKNI_YEAST	P33336 saccharomyc
26	94.5	4.2	283	1 BRUI_SOYBN	P35694 glycine max
27	93.5	4.2	1052	1 MSIP_SCHPO	O92288 cricetus
28	93.5	4.2	2397	1 MORE_SCHPO	O09854 schizosacch
29	92.5	4.1	380	1 TREB_MYCTU	O10769 mycobacteri
30	92.5	4.1	931	1 MAND_RAT	O9un23 aspergillus
31	92.5	4.1	1052	1 MSIP_RAP	O9wt33 ratius norv
32	91	4.1	550	1 MYCO_STRCI	P20910 streptomyc
33	91	4.1	995	1 AGAA_VIBS7	P48839 vibrio sp.

34	90.5	4.0	534	1 FM2_ACTNA	P12616 actinomyces
35	90	4.0	431	1 KRE2_CANAL	O00310 candida alb
36	89.5	4.0	309	1 AGAR_STRCO	P07883 streptomyc
37	89.5	4.0	491	1 FIBB_HUMAN	P02675 homo sapien
38	89.5	4.0	537	1 APE3_YEAST	P37302 saccharomyc
39	89.5	4.0	541	1 COX1_BRAUA	P31833 bradyrhizob
40	89	4.0	259	1 GUB_BACBR	P37073 bacillus br
41	89	4.0	776	1 ISOA_PSEAY	P10342 pseudomonas
42	89	4.0	776	1 ISOA_PSEAP	P26501 pseudomonas
43	88.5	4.0	351	1 DESA_SPIPL	O54794 spirulina p
44	88.5	4.0	779	1 ACON_GRAVE	P49609 giacellaria
45	88.5	4.0	900	1 GUNH_CLOTH	P16218 clostridium

ALIGNMENTS

RESULT 1

ID	E13B_BACCI	STANDARD:	PRT:	682 AA.
AC	P23903;			
DT	01-NOV-1991 (Rel. 20, created)			
DT	01-NOV-1991 (Rel. 20, last sequence update)			
DT	15-JUL-1999 (Rel. 38, last annotation update)			
DE	Glucan endo-1,3-beta-glucosidase A1 precursor (EC 3.2.1.39) ((1->3)-beta-glucan endohydrolase) ((1->3)-beta-glucanase A1).			
GN	GLCA.			
OS	Bacillus circulans.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxID=1397;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 39-52.			
RC	STRAIN-WL-12;			
RA	MEDLINE=90185240; PubMed=2311931;			
RA	Yabata N., Matanabe T., Nakamura Y., Yamamoto Y., Kamimura S., Tanaka H.;			
RT	"Structure of the gene encoding beta-1,3-glucanase A1 of Bacillus circulans WL-12."			
RL	Gene 86:113-117(1990).			
CC	- FUNCTION: LYSIS OF CELLULAR WALLS CONTAINING BETA-1,3-GLUCANS.			
CC	- IMPLICATED IN THE DEFENSE AGAINST FUNGAL PATHOGENS.			
CC	- CATALYTIC ACTIVITY: Hydrolysis of 1,3-beta-D-glucosidic linkages in 1,3-beta-D-glucans.			
CC	- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib.ch).			
CC	EMBL: M34503; AAA22474.1; -			
DR	PIR: J00420; J00420.			
DR	InterPro: IPR000757; Glyco_hydro_16.			
DR	Pfam: PF00722; Glyco_hydro_16; 1.			
DR	PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.			
KW	Cell wall; Hydrolase; Glycosidase; Signal.			
FT	SIGNAL			
FT	CHAIN			
FT	ACT_SITE			
FT	ACT_SITE			
SO	SEQUENCE			

Query Match 15.7%; Score 351.5; DB 1; Length 682;
Best Local Similarity 26.2%; Pred. No. 2.2e-21;
Matches 100; Conservative 44; Mismatches 89; Indels 149; Gaps 14;

OY 25 HIWODEFD-YFDGAKQHEV-----TATGGNSEPOLYQDGNANSVVRGKLFK 74
DB 424 NLIMWDEFNGTTIDTSKMYETGYLLNNDPATWGMGNAELGYHTNSTONVYVQDGKLIK 483

[illegible]

RESULT 2			
GUB_RHOMR			
AC	GUB_RHOMR	STANDARD:	PRT: 286 AA.
AC	P45798:		
DT	01-NOV-1995 (Rel. 32, Created)		
DT	01-NOV-1995 (Rel. 32, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)		
DE	(1,3-1,4-Beta-D-glucan 4-glucanohydrolase) (Lichenase).		
GN	BGLA.		
OS	Rhodothermus marinus (Rhodothermus obamensis).		
OC	Bacteria; Bacteroidetes; Sphingobacteriales; Sphingobacteriales;		
OC	Crenitrichaceae; Rhodothermus.		
OX	NCBI_TaxID=29549;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RP	STRAIN=21 / ITI-378;		
RA	MEDLINE=95010084; PubMed=7925416;		
RA	Spilliaert R., Hreggvidsson G.O., Kristjansson J.K.,		
RA	Egertersson G., Palsdottir A.;		
RT	*Cloning and sequencing of a Rhodothermus marinus gene, bglA, coding		
RT	for a thermostable beta-glucanase and its expression in Escherichia		
RT	coli.;		
RL	Eur. J. Biochem. 224:923-930(1994).		
CC	-1- FUNCTION: SHOWS ACTIVITY ON LICHENAN, BETA-GLUCAN AND LAMINARIN		
CC	BUT NOT ON CMC CELLULOSE OR XILAN.		
CC	-1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages		
CC	in beta-D-glucans containing 1,3- and 1,4-bonds.		
CC	-1- MISCELLANEOUS: The enzyme has a temperature optimum of 85 degrees		
CC	Celsius and a pH optimum of 7.0.		
CC	-1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on lists		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@sib-sib.ch).		
CC	-----		
DR	EMBL: U04836; AAA60459.1; .		
DR	PIR: S48201; S48201.		
DR	InterPro: IPR000757; Glyco_hydro_16.		
DR	Pfam: PF00722; Glyco_hydro_16; 1.		
DR	PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.		
KW	Hydrolase; Glycosidase; Signal.		
FT	SIGNAL	1	POTENTIAL.
FT	CHAIN	31	286
FT			BETA-GLUCANASE.

FT	ACT_SITE	158	158	NUCLEOPHILE (BY SIMILARITY).
PT	ACT_SITE	163	163	PROTON DONOR (BY SIMILARITY).
SQ	SEQUENCE	286 AA:	33145 MW;	7215C33624135191 CRC64;

Query Match	15.4%	Score 344.5	DB 1	Length 286
Best Local Similarity	27.5%	Pred. No. 3.1e-21		
Matches 112; Conservative	37	Mismatches 96	Indels 163	Gaps 17

```

Oy      1 MRRLVLVCLLFEGEFAFTD-----WDOYIIVODEDYF-----DGAKNOHEVTAIGG 50
Db      13 MRRIAFILSLVILGCMGSDRDKAPHF--LVMSDEPDYSGLPDPKMDVDYGGHGNG 69
Oy      51 NSEFOLYQOD-ANSEVFDGKLFIRPTLLADNINPQTGAFCSTDEMYNGVLDMVAMYGAC 109
Db      70 NQELQYTRARLENAKRVGGVLLIEA-----RHEPEGREY-----105
Oy      110 TINTDNNCGYRTGACADIPRMSARVTRTOKYSFTIGRUVVYHAKMPRGDMLPRAIMLPED 169
Db      106 -----TSARLVTRKASMTYGRFIRARLPBGRGTWPAIMLIPDR 145
Oy      170 WYVYG--WPRSGEIDIIETIG-NRD-----FKNNGGEFLGLOKMSITHMGPGMDNRY 220
Db      146 QTYGSAYTPDNGEIDIMEHGFNPDPVHGCTVHTKAYINHLGTORGGS-----192
Oy      221 WLTSLPKSHSDMNYNDNFHTEWFDSPNGLRFREVD-----NQALIDVPYPLIDANPW 274
Db      193 --INVTAKRD-----PHYAIEMTPEIRIRFVDSLYRRFPNEULTD---PEADMRHW 241
Oy      275 WVDWFEMWGPMLPOYENDNPMWAGTGNLAPFDONFHEILVAVAGTNGEITPDGGINRGDP 334
Db      242 -----PFDQPFHLINNIAGVGAWGG-----OQGVDP 267
Oy      335 ALQRPWSNGDWINDAMKRFDPARGMKMTWTDDEGDNANAGYIIRYRK 382
Db      268 -----EAFPAQ-----LVVDIVRYR 283

```

RESULT	3			
EXSH_RHIME				
ID	EXSH_RHIME	STANDARD;	PRT;	465 AA.
AC	033680;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Endo-1,3-1,4-beta-glucanase exsh (EC 3.2.1.-) (Succinoglycan			
DE	biosynthesis protein:exsh).			
GN	EXSH OR RB1055 OR SMB2093Z.			
OS	Rhizobium meliloti (Sinorhizobium meliloti).			
OS	Plasmid pSYM (megaplasmid 2).			
CC	Bacteriae; Proteobacteria; Alphaproteobacteria; Rhizobiales;			
CC	Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.			
CC	NCBI_TaxID=382;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21899394; PubMed=11902715;			
RA	York G.M., Walker G.C.;			
RT	"The Rhizobium meliloti exoX gene and prsp/prse/exsh genes encode			
RT	components of independent degradative pathways which contribute to			
RT	production of low-molecular-weight succinoglycan.";			
RL	Mol. Microbiol. 25:117-134(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=1021;			
RX	MEDLINE=21396508; PubMed=11481431;			
RA	Fitnan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,			
RA	Vorhoefer F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,			
RA	Goulding B., Puehler A.;			
RT	"The complete sequence of the 1,683-kb pSYM megaplasmid from the N2			
RT	fixing endosymbiont Sinorhizobium meliloti.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).			
RN	[3]			
RP	CHARACTERIZATION,			
RX	MEDLINE=98226741; PubMed=9560202;			

RA York G.M., Walker G.C.;
 RT The Rhizobium meliloti ExoX and ExoH glycanases specifically
 RL depolymerize nascent succinoglycan chains.;
 CC Proc. Natl. Acad. Sci. U.S.A. 95:4912-4917(1998).
 CC -1- FUNCTION: CLEAVES HIGH MOLECULAR WEIGHT SUCCINOGLYCAN TO YIELD LMW
 CC SUCCINOGLYCAN. DYNAMICALLY REGULATE THE MOLECULAR WEIGHT
 CC DISTRIBUTION OF SUCCINOGLYCAN BY CLEAVING NASCENT SUCCINOGLYCAN
 CC ONLY DURING A LIMITED PERIOD AFTER ITS SYNTHESIS. PERHAPS BEFORE
 CC IT UNDERGOES A TIME-DEPENDENT CHANGE IN ITS CONFORMATION OR
 CC AGGREGATION STATE.
 CC -1- PATHWAY: Exopolysaccharide biosynthesis.
 CC -1- SUBCELLULAR LOCATION: SECRETED. PROBABLY BY A TYPE-III SECRETION
 CC SYSTEM.
 CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U89164; AAB64093.1; -
 DR EMBL: AL603645; CAC49455.1; -
 DR PIR: G95973; G95973.
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR InterPro: IPR001343; Hemolysn_Ca_bind.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 DR Pfam: PF00353; hemolysin_cabind; 3.
 DR PRINTS: PR00133; CABNDNGRPT.
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; FALSE_NEG.
 DR PROSITE: PS00330; HEMOLYSIN_CALCITUM; FALSE_NEG.
 DR Exopolysaccharide synthesis; Glycosidase; Hydrolase; Plasmid;
 KW Complete proteome.
 FT DOMAIN 275 465 CATALYTIC.
 FT ACT_SITE 349 349 NUCLEOPHILE (BY SIMILARITY).
 FT ACT_SITE 354 354 PROTON DONOR (BY SIMILARITY).
 FT SEQUENCE 465 AA; 50286 MW; 6C848236E9E8C8A CRC64;
 SO
 Query Match 7.8%; Score 175.5; DB 1; Length 465;
 Best Local Similarity 21.6%; Pred. No. 4e-07;
 Matches 77; Conservative 40; Mismatches 102; Indels 137; Gaps 15;
 QY 27 VNQDEDFYDGAQKQHEVTATGGSEFQLYTDGANSFVRDGLKPLKTLADNINPQT 86
 DB 238 VWDARF-----WMAPEKGAATLSSNGEQWYINPS-----YEPFA---SVNPF 277
 QY 87 GAFPGDFMYNGVLQVWAMYGA-CTNTDNNGCYRTGAAGDIPPMASARVETPKYSTFHG 145
 DB 278 -----VNVGVLTITTAAPASFAIDAEINGDYDT-----SGMLTYSSFAQTYG 319
 QY 146 RYVVAHAKPVGDMIAWMLPEDWYVGGPRSGEIDITIGNRDKFGKNGEFLGQKM 205
 DB 320 YEMRADMPDDGQVMAFWLLPAD---GSRP--PELDVYEMKQD---SNT-----Y 363
 QY 206 GSTMHWGPGMDNRKWLSTLPHKSDWNTGDNHTFTWFWPMSPGNGLEFVDDENQALDVP 265
 DB 364 IATVH--SNETGSRSTSIENSVAAD---ASGEHTYGVLMTEBEIYWFYFDADAIRADTP 417
 QY 266 YLLIDANFWVDFWENGKRWLPQYENDNFWAGSTNLAPFQONHFLNLNVAVGSTNGFIPD 325
 DB 418 SDMHQ-----PMT-----MNVNLAAGGAGTAPRD 441
 QY 326 GCINRGDPAIQPMNSGDWYNDAMRKFFDARGMFKMTWDEGDNNAMQDYIRYV 381
 DB 442 GL-----ADGSEMKIDITIKAY 457
 RESULT 4
 EGIQ_RHIME STANDARD; PRT; 465 AA.
 ID EGIQ_RHIME
 AC Q92302;

DT 30-MAY-2000 (Rel. 39, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Endo-1,3-1,4-beta-glycanase egIC (EC 3.2.1.-) (Succinoglycan
 DE biosynthesis protein egIC).
 GN EGIQ OR RA0864 OR SWA1587
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Plasmid psyma (megaplasmid 1).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 OX NCBI_TaxID=382;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021.
 RX MEDLINE=21396509; PubMed=10485295;
 RA Sharypova L.A., Yurgel S.N., Keller M., Simerov B.V., Puehler A.,
 RA Becker A.;
 RT "The eff-482 locus of Sinorhizobium meliloti CML-105 that influences
 RT symbiotic effectiveness consists of three genes encoding an
 RT endoglycanase, a transcriptional regulator and an adenylate cyclase.";
 RL Mol. Gen. Genet. 261:1032-1044(1999).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021.
 RX MEDLINE=21396509; PubMed=11481432;
 RA Barnett M.J., Fisher R.F., Jones T., Komp C., Aboja A.P.,
 RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
 RA Gural J.M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
 RA Khatman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
 RA Yeh K.-C., Davis R.W., Federpsiel N.A., Long S.R.;
 RT "Nucleotide sequence and predicted functions of the entire
 RT Sinorhizobium meliloti psyma megaplasmid.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
 CC -1- FUNCTION: CLEAVES HIGH MOLECULAR WEIGHT SUCCINOGLYCAN TO YIELD LMW
 CC SUCCINOGLYCAN. DYNAMICALLY REGULATE THE MOLECULAR WEIGHT
 CC DISTRIBUTION OF SUCCINOGLYCAN BY CLEAVING NASCENT SUCCINOGLYCAN
 CC ONLY DURING A LIMITED PERIOD AFTER ITS SYNTHESIS. PERHAPS BEFORE
 CC IT UNDERGOES A TIME-DEPENDENT CHANGE IN ITS CONFORMATION OR
 CC AGGREGATION STATE (BY SIMILARITY).
 CC -1- PATHWAY: Exopolysaccharide biosynthesis.
 CC -1- SUBCELLULAR LOCATION: SECRETED. PROBABLY BY A TYPE-III SECRETION
 CC SYSTEM (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AJ225896; CAB38101.1; -
 DR EMBL: AB002723; AAR65522.1; -
 DR PIR: H93369; H93369.
 DR InterPro: IPR000757; Glyco_hydro_16.
 DR InterPro: IPR001343; Hemolysn_Ca_bind.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 DR Pfam: PF00353; hemolysin_cabind; 3.
 DR PRINTS: PR00133; CABNDNGRPT.
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; FALSE_NEG.
 DR PROSITE: PS00330; HEMOLYSIN_CALCITUM; FALSE_NEG.
 DR Exopolysaccharide synthesis; Glycosidase; Hydrolase; Plasmid;
 KW Complete proteome.
 FT DOMAIN 275 465 CATALYTIC.
 FT ACT_SITE 349 349 NUCLEOPHILE (BY SIMILARITY).
 FT ACT_SITE 354 354 PROTON DONOR (BY SIMILARITY).
 FT CONFLICT 52 52 I -> T (IN REF. 1).
 FT SEQUENCE 465 AA; 49614 MW; 12CB879AED9E6558A CRC64;
 SO
 Query Match 7.5%; Score 167; DB 1; Length 465;
 Best Local Similarity 22.7%; Pred. No. 2e-06;
 Matches 82; Conservative 40; Mismatches 117; Indels 122; Gaps 19;

```

QY 16 PAFDMDQYH-----VMDPEFYF---DGAK-----MOHEVTATGGG 50
DB 196 FANTSAQOLHANGPSIALDRSVLQOTPSDDPNTIQLSGTSGVMDPKWMAPEKATILTG 255
QY 51 NSEFQLYTQDANSFVBDGKLEIKPPLADNINPOTGA-PRGTFMNGVLDMVAMYGA- 108
DB 256 NDELQWY-----VNP-----YOPTASANPS-----VTDGLVLTITAKPAQS 292
QY 109 CTNNNDNGCYRTGAGDIPRMSARVRFQKXSTFHGVVYVHAKKPVCDMLPRAIMLPE 168
DB 293 AIAQETNGYDT-----SGMLTYSSEFAQTYGFEKRAMDPDQAGAMPAPWLLPG 342
QY 169 DMVYGCWPRSGEIDIIETIGNRDKNGGEFLGLOKMSHTMNGWPGMDNRYWLTSLPKH 228
DB 343 D---GTPP---PELDVEMHGD---PNYIATVHSHNGSOTS-----TASARV 384
QY 229 SDDMNNGDNHRTFEMFNSPGLREFVDENQALLDVPYPLIDANFMYVDFEMWCKPWLPO 288
DB 385 TDT-----SGFHKYGVLTWEETIVWYFDDAIARADTSPDMHD-----PM 424
QY 289 YENDNPAAGTNLAPPONFILNVAVGNGEFLPDGCTNRGGDPAIQ--KPMNS-NDWM 345
DB 425 Y-----MLVNLAIAGNAGPPTDLMG-GAEKVDYKAYSLDWM 463
QY 346 Y 346
DB 464 H 464

```

RESULT 5
DHET ACEPO STANDARD: PRT: 738 AA.

AC P28036:
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
ADHA.
OS Acetobacter polyoxogenes.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Acetobacteraceae; Acetobacter.
OX NCBI_TaxID=439;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN-NB11028;
RX MEDLINE-91159482; PubMed-2001402;
RA Tamaki T., Fukaya M., Takemura H., Tayama K., Okumura H.,
Kawamura Y., Nishiyama M., Horinouchi S., Beppu T.;
*Cloning and sequencing of the gene cluster encoding two subunits of
membrane-bound alcohol dehydrogenase from acetobacter polyoxogenes.;
RT Blochum. Biophys. Acta 1088:292-300(1991).
RL Blochum. Biophys. Acta 1088:292-300(1991).
CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: POO AND HEME.
CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
CC COMPLEX.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
CC SPACE (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC -1- SIMILARITY: Contains 1 cytochrome c domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D00635; BAA00528.1; -

```

DR PIR; S14270; S14270.
DR HSSP; 0924J7; 1FEG.
DR InterPro; IPR001479; Bac_POO.
DR InterPro; IPR002372; Bac_POO_repeat.
DR InterPro; IPR000345; CytC_heme_bind.
DR Pfam; PF01011; Bacterial_POO; 3.
DR SMART; SM00564; POO; 3.
DR PROSITE; PS00363; BACTERIAL_POO_1; 1.
DR PROSITE; PS00364; BACTERIAL_POO_2; 1.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
DR Oxidoreductase; POO; Heme; Periplasmic; Membrane; Signal.
KM SIGNAL 1 35
FT CHAIN 36 738
FT DOMAIN 634 738
FT BINDING 650 738
FT BINDING 653 653
FT BINDING 654 654
FT METAL 654
SQ SEQUENCE 738 AA; 80840 MW; 1E2B6ED7BCD92AF6 CRC64;

Query Match 5.0%; Score 112.5; DB 1; Length 738;
Best Local Similarity 20.4%; Pred. No. 0.098;
Matches 78; Conservative 37; Mismatches 113; Indels 155; Gaps 21;

QY 48 GGNSEFQLYTQDANSFVR-----DGKLFKPTLLADNINPOTGAPGTD-FMYN 97
DB 212 GNGGSEF-----GARGFVSAPDAETGKYDMRFETVP-----NPKNPPDAASDVLNN 258
QY 98 GVLDMWAMYGACTNTDNNCC-----VETGAAGD----- 125
DB 259 KAYQTFSPGAWTRGGGTWDSIVYDPAVLVYLVGNGSPWNYKRYSEGGKDLFLTG 318
QY 126 ---IPRMSARVRFQ-----KYSFTBGR-----VVVHAK----- 152
DB 319 STVALKPEYGEYVWFQETPRHQMDFTSQOIMTLDLPINGEIRHIVARKNGFYIID 378
QY 153 MFVGDMLPAMLPEDWYVGGW-----PRSEIDIIETIGNRD--FKNTGGEFLGIQ-K 204
DB 379 AKTGEFI-----SGKNVYVWNASGLDPKGTG----RPIYNDADLYLTGKEMVGIPGD 427
QY 205 MG-----STMHWGPGMDNDNRVYLTSLP-----KHSDDMYGDNFHFWD 244
DB 428 LGHNFPAAVAFSP-----KTGLVYIPAQVPLTYNQGVGFPHPHSMNLGLDMNVVGI 482
QY 245 WSPNGLRFFVDENQALLDVPYPLIDANFMYVDFWE---WGRFWLRYENDNFWAGTNNL 301
DB 483 DSPEAKQAFVRLK-----GVIAMDPOQAQAEW---RVHKKPGWNGGILA 525
QY 302 APFDONFHEILN---VAVGTNG 321
DB 526 TGGDLLFOGLANGFHFAYDATNG 548

```

RESULT 6
XVNA_RUMFL
ID XVNA_RUMFL STANDARD: PRT: 954 AA.
AC P29126;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bifunctional endo-1,4-beta-xylanase Xyla precursor (EC 3.2.1.8).
GN XVNA.
OS Ruminooccus flavefaciens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae;
OC Ruminooccus.
OX NCBI_TaxID=1265;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-17;
RX MEDLINE-92261318; PubMed-1584021;
RA Zhang J.-X., Flint H.J.;
RT "A bifunctional xylanase encoded by the xyna gene of the rumen
RT cellulolytic bacterium Ruminooccus flavefaciens 17 comprises two
RT dissimilar domains linked by an asparagine/glutamine-rich sequence.";

DR PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
 KW Hydrolyase; glycosidase; signal; 3D-structure.
 FT SIGNAL 1 27
 FT CHAIN 28 243
 FT ACT_SITE 134 134
 FT ACT_SITE 138 138
 FT ACT_SITE 138 138
 FT DISULFID 61 90
 FT MUTAGEN 51 51
 FT MUTAGEN 89 89
 FT MUTAGEN 92 92
 FT MUTAGEN 105 105
 FT MUTAGEN 133 133
 FT MUTAGEN 134 134
 FT MUTAGEN 136 136
 FT MUTAGEN 138 138
 FT MUTAGEN 143 143
 FT MUTAGEN 160 160
 FT MUTAGEN 168 168
 FT MUTAGEN 179 179
 FT MUTAGEN 190 190
 FT MUTAGEN 219 219
 FT STRAND 35 37
 FT STRAND 44 46
 FT STRAND 47 49
 FT STRAND 53 53
 FT STRAND 57 58
 FT STRAND 59 59
 FT STRAND 62 63
 FT STRAND 65 67
 FT HELIX 68 70
 FT STRAND 72 73
 FT STRAND 76 84
 FT STRAND 85 86
 FT STRAND 87 95
 FT STRAND 99 99
 FT STRAND 102 109
 FT STRAND 114 115
 FT STRAND 116 124
 FT STRAND 126 128
 FT HELIX 129 129
 FT STRAND 133 140
 FT STRAND 141 142
 FT STRAND 144 145
 FT STRAND 146 153
 FT STRAND 154 155
 FT STRAND 156 156
 FT STRAND 161 164
 FT STRAND 169 171
 FT STRAND 174 181
 FT STRAND 182 183
 FT STRAND 184 189
 FT STRAND 190 191
 FT STRAND 192 197
 FT STRAND 206 215
 FT STRAND 219 222
 FT HELIX 231 243
 FT STRAND 231 243
 SO SEQUENCE 243 AA; 27435 MW; 651188D9AAD609A5 CRC64;

Query Match 4.9%; Score 110; DB 1; Length 243;
 Best Local Similarity 21.9%; Pred. No. 0.046;
 Matches 60; Conservative 42; Mismatches 88; Indels 84; Gaps 18;

QY 71 LFIKRLADININPTGAPFGDF-MYNGVLDVWVWYACCTDNNCGYRGAAGDIPPA 129
 DB 16 LFLISTRAASASAOGTGSGFYEPNNYN--TGLMKKADGYSNGNMFNC--TKRANNVSKT 71
 QY 130 MSARVR-----TFQKY-----SFTHGRRVVVHAKKPVQDWLMPA--IMLPEDWVY 172
 DB 72 SLGKRLSLTSPSYKPKFCGGRNSVQTYGYLVEVNMK-----PAKNVGIVSSFTY 123
 QY 173 GGPWPSG---EIDILEITGNDFKNTGGEPLGIGIKMGSTYMWGPGMDNRYWLTSLPKH 228
 DB 124 TG-PTDGTWDEIDI-----EFLG--KDTTKVQF-----NYTNGVGNH 159

QY 229 SDWMYV-----DNFHTFEMDMSPNGLRFFVDE--NOALLDVPY--LIDANPW---WVD 277
 DB 160 EKIVNLGFGPAANSYHYAFADWQPNISKWVVDQGLKRTATQTQIPQGTGKIMMNMNAGVD 219
 QY 278 FWEWGPWLPOYENDPNWAGTTLAPEDONFHEI 311
 DB 220 -----EWLGSY-----NGVTPLSRSLHW 238

RESULT 8
 ID NANH_VIBCH STANDARD; PRT; 781 AA.
 AC P37060; O9KR59;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Stalidase precursor (EC 3.2.1.18) (Neuraminidase) (NANase).
 GN NANH OR VC1784.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_Taxid=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eissen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermlaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae.";
 RL Nature 406:477-483(2000).
 RN [3]
 RP SEQUENCE OF 1-44 FROM N.A., AND SEQUENCE OF 25-44.
 RC STRAIN=Classical Ogawa 395 / ATCC 39541 / Serotype O1;
 RX MEDLINE=88169467; PubMed=2832365;
 RA Virm E.R., Lawlisuk L., Galen J.E., Kaper J.B.;
 RT "Cloning and expression of the Vibrio cholerae neuraminidase gene
 RT nanB in Escherichia coli.";
 RL J. Bacteriol. 170:1495-1504(1988).
 RN [4]
 RP CHARACTERIZATION, AND X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=92389334; PubMed=1518058;
 RA Taylor G.L., Virm E.R., Garman E.F., Laver W.G.;
 RT "Purification, crystallization and preliminary crystallographic study
 RT of neuraminidase from Vibrio cholerae and Salmonella typhimurium
 RT LT2.";
 RL J. Mol. Biol. 226:1287-1290(1992).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE=95006320; PubMed=7922030;
 RA Creneils S., Garman E.F., Laver W.G., Virm E.R., Taylor G.L.;
 RT "Crystal structure of Vibrio cholerae neuraminidase reveals dual
 RT lectin-like domains in addition to the catalytic domain.";
 RL Structure 2:535-544(1994).
 CC -I- FUNCTION: CLEAVES THE TERMINAL SIALIC ACID (N-ACETYL NEURAMINIC
 CC ACID) FROM CARBOHYDRATE CHAINS IN GLYCOPROTEINS PROVIDING FREE
 CC SIALIC ACID WHICH CAN BE USED AS CARBON AND ENERGY SOURCES.
 CC SIALIDASES HAVE BEEN SUGGESTED TO BE PATHOGENIC FACTORS IN

[illegible]

FT	STRAND	224	226
FT	TURN	232	236
FT	STRAND	238	246
FT	TURN	247	247
FT	TURN	251	252
FT	TURN	254	255
FT	STRAND	257	265
FT	TURN	266	267
FT	STRAND	276	277
FT	TURN	280	282
FT	STRAND	288	298
FT	TURN	299	302
FT	STRAND	303	312
FT	TURN	313	314
FT	TURN	317	317
FT	HELIX	318	320
FT	TURN	324	325
FT	STRAND	339	336
FT	TURN	337	340
FT	STRAND	341	347
FT	HELIX	349	352
FT	STRAND	357	360
FT	TURN	362	363
FT	STRAND	364	371
FT	TURN	376	377
FT	STRAND	380	391
FT	STRAND	404	400
FT	STRAND	403	411
FT	TURN	413	414
FT	STRAND	417	421
FT	STRAND	428	431
FT	HELIX	434	437
FT	TURN	438	438
FT	STRAND	441	448
FT	TURN	449	452
FT	STRAND	453	458
FT	TURN	459	460
FT	STRAND	461	466
FT	STRAND	470	470
FT	STRAND	474	480
FT	STRAND	487	498
FT	TURN	499	500
FT	STRAND	501	507
FT	HELIX	508	512
FT	TURN	513	513
FT	HELIX	516	517
FT	HELIX	522	525
FT	TURN	526	526
FT	STRAND	528	533
FT	STRAND	536	539
FT	STRAND	543	545
FT	STRAND	552	553
FT	TURN	558	559
FT	TURN	561	562
FT	TURN	565	566
FT	STRAND	568	575
FT	STRAND	580	587
FT	TURN	590	591
FT	STRAND	595	599
FT	STRAND	605	608
FT	TURN	609	610
FT	STRAND	611	613
FT	STRAND	617	624
FT	TURN	626	627
FT	STRAND	630	636
FT	STRAND	641	642

Query Match 4.9%; Score 109; DB 1; Length 781;
Best Local Similarity 23.6%; Pred. No. 0.2;
Matches 82; Conservative 31; Mismatches 107; Indels 128; Gaps 20;


```

OY 324 -----PDGCRNGDPAALOKPWS-----NGDM-VNDAMRKFPDARGMKTWDEGDN 370
DB 606 AKALHPDG--NIGMRIRISKEPMSIILNLCISNMWAIIDWOYIFPV-----649
OY 371 NAMOVYIRVYK 382
DB 650 -VMSIDYIRIYO 660

RESULT 10
ID GUB_PAEPO STANDARD: PRT: 238 AA.
AC P45797:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
GN (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).
OS Paenibacillus polymyxa (Bacillus polymyxa).
OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Paenibacillus.
ON NCBI_TaxID=1406;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 842:
RX MEDLINE-92041687; PubMed-1938968;
RA Gosalbes M.J., Perez-Gonzalez J.A., Gonzalez R., Navarro A.;
RT "Two beta-glucanase genes are clustered in Bacillus polymyxa:
RT molecular cloning, expression, and sequence analysis of genes
RT encoding a xylosidase and an endo-beta-(1,3)-(1,4)-glucanase.";
RL J. Bacteriol. 173:7705-7710(1991).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X57094; CAA40379.1; -
DR PIR: S19012; S19012.
DR HSSP: P23904; ICPN.
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00722; Glyco_hydro_16.
DR PRINTS: PR00737; GLHYDRASE16.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
KW Hydrolase; Glycosidase; Signal.
FT SIGNAL 1 26
FT CHAIN 27 238
FT ACET_SITE 129 129
FT ACET_SITE 133 133
FT DISULFID 56 85
FT SEQUENCE 238 AA; 26919 MW; CQCF784EASD40E8C CRC64;

Query Match 4.7%; Score 106; DB 1; Length 238;
Best Local Similarity 22.0%; Pred. No. 0.095;
Matches 62; Conservative 32; Mismatches 92; Indels 96; Gaps 16;

OY 25 HIVMODEEDYDGAKEHVEVATGGNSEFOLYTODGAN-SFVRGKLEIKETLLADNIN 83
DB 28 NVEW-EPLSYENSSWQ--KADGSGNQMPCNCTRRANVNTNNGKLSLTSFANN-- 81
OY 84 PGTGAPGTDFMTNGLDVWMTYAGACTNTDNGCYRTGAAGDIPRMSARVTFQKYSF- 142
DB 82 -----KFDG-----GEYRSTNNYG-YGLVEVSMKPAKNGTIVSFFYITGP 121
OY 143 THGRVYVAKMPVGMPLPALMLPDMVYGGWPRSGEIDITETIGNRDFKNTGGEFLGI 202

```

```

DB 122 SHG-----TCW-----DEID-----EFLG- 136
OY 203 QKMGSTMHGPGWDDNRMYLTLSPKHSDDNNGY-----DNHTEFWMSPNGLREFVDD- 256
DB 137 -KDTTKVQF-----NYITNGVGGEHKIITNLGFDASTSHTYAFDQPGIKRYVGVYL 188
OY 257 ENQALLDVPYPLIDANPMWVDFEWG---KPMILOYENDNP 294
DB 189 KHTATNIP-----STPGKIMNLMNMGTVDSWLGSYNGANP 225

RESULT 11
ID DHEI_ACEU STANDARD: PRT: 739 AA.
AC Q44002; Q07952;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alcohol dehydrogenase (acceptor) precursor (EC 1.1.99.8).
GN ADH.
OS Acetobacter europaeus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Acetobacteraceae; Gluconacetobacter.
ON NCBI_TaxID=33995;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DESI / DSM 6160;
RA Thurner C.A.K.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor -> an aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: POQ AND HEME (BY SIMILARITY).
CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
CC COMPLEX (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
CC SPACE (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -1- SIMILARITY: Contains 1 cytochrome c domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X82894; CAA58066.1; -
DR EMBL: Y09480; CAA70688.1; -
DR HSSP: Q924J7; IFLG.
DR InterPro: IPR001479; Bac_POQ.
DR InterPro: IPR002372; Bac_POQ_repeat.
DR InterPro: IPR000345; CytC_heme_bind.
DR Pfam: PF01011; Bacterial_POQ; 3.
DR SMART: SM00564; POQ; 3.
DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
DR PROSITE: PS00190; CYTOCHROME_C; 1.
KW Oxidoreductase; POQ; Heme; Feridplasmic; Membrane; Signal.
FT SIGNAL 1 35
FT CHAIN 36 739
FT DOMAIN 635 739
FT BINDING 651 651
FT BINDING 654 654
FT METAL 655 655
FT SEQUENCE 739 AA; 80944 MW; E681BB37ACB91F4 CRC64;

Query Match 4.7%; Score 106; DB 1; Length 739;
Best Local Similarity 20.6%; Pred. No. 0.34;
Matches 78; Conservative 37; Mismatches 118; Indels 146; Gaps 20;

```

QY 48 GGNSEFQLYTQDGNSEFV-----DGKLFKPTLLADNINP----- 84
 DB 212 GNGSEF-----GARGFVTAFDAETGKVDNREFAP-----NPKNEPDHTASDVLM 258
 QY 85 -----QTAPFCT-----DFATNGVLDAVAMGACINIDNNGCYRTGAAGD--- 125
 DB 259 NKAVQTSPTGAMTRQGGGTVWDSIVYDPAVDL-VYLGVGNGSPWNRKRYREGKDNLF 317
 QY 126 -----IPPAMSAVRFPQ-----KYSFTHGRVYVHAKPMGDMWPAIMMLP----- 167
 DB 318 LGSIVALKPTEGVEYVNHQETPMDQMDTTSVOQITLPLPINGEYRHHVIAHAKPKGFFYI 377
 QY 168 -----EDWYGGW-----PRSGEIDIETIGNRD--FKNTGSEFLGIQ-KMG-- 206
 DB 378 IDAKTGEFISGKNVYVVMASGLDPKGTG-----RTYMPDALYTLTGKEMWGIPIGDLGGH 432
 QY 207 --STMHGPGMDNRWLTSLP-----KHSDDNRYGNGFHTFWDSPN 248
 DB 433 NFAMAFSP-----KTGLVYIPAOQVPELYTNOVGGEFPHPSWMLGIDMNVKGIPIDSP 487
 QY 249 GLREFVDENQALLDVPRPLIDANPMWDFWE--WCKPWLPOYENDPMAGTMLAFPD 305
 DB 488 AKQAFVADLK-----GWLVAWDPKQKQAAW--RVDHKGPMNGGIIATGSD 530
 QY 306 QNEFFILN--VAVGGTNG 321
 DB 531 LLEOGLANGEFHAYDATING 549

RESULT 12
 OSTA_HAEIN STANDARD; PRT; 782 AA.
 AC P44846;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Organic solvent tolerance protein precursor.
 GN IMP OR OSTA OR H10730.
 OS Haemophilus influenzae.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 CC Pasteurellaceae; Haemophilus.
 OC NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Rd / KM20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7543800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McInerney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uitterlbeck T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd.";
 RL Science 269:496-512(1995).
 RN [2]
 RP IDENTIFICATION BY MASS SPECTROMETRY.
 RX MEDLINE=20137488; PubMed=10675023;
 RA Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
 RA Gray C., Fountoulakis M.;
 RT "Two-dimensional map of the proteome of Haemophilus influenzae.";
 RL Electrophoresis 21:411-429(2000).
 CC -I- FUNCTION: Determines N-hexane tolerance. Involved in outer
 CC membrane permeability. Essential for envelope biogenesis. Could be
 CC part of a targeting/usher system for outer membrane components (by
 CC similarity).
 CC -I- SUBCELLULAR LOCATION: Outer membrane (by similarity).
 CC -I- SIMILARITY: BELONGS TO THE IMP/OSTA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U32756; AAC2389.1; -
 DR PIR: G64157; G64157.
 DR TIGR: H10730; -
 DR HAMAP: MF_01411; -; 1.
 DR InterPro: IPR005653; Osta.
 DR Pfam: PF03968; Osta; 1.
 DR Pfam: PF04453; Osta.C; 1.
 KW Outer membrane; Signal; Complete proteome.
 FT SIGNAL 1
 FT CHAIN 24
 SQ SEQUENCE 782 AA; 90084 MW; E73C8A578B02D1B CRC64;

Query Match 4.7%; Score 106; DB 1; Length 782;
 Best Local Similarity 21.3%; Pred. NO. 0.36; Mismatches 112; Indels 148; Gaps 26;
 Matches 83; Conservative 46;

QY 32 FDYFDGAKWHEVTATGGGSEFQLYTQDG--ANS--FV-RDGKLFKPTLLADNINPQ 85
 DB 106 FDYKD-----NOINML-GKDAEFNLDSDHGNLTNSEYEVGROGR-----GKADNIT-- 151
 QY 86 TGAPFGTDPMYNGVLDAVAMGACINTDNNCGYRTGAAGDIPAMSA-RVRFQKYSFT- 143
 DB 152 -----LHNN-----YRVMKNATFTSLH-----GDNMAVADASEIRQYKKEVAE 191
 QY 144 --HGRVYVHA-----KMPVG-----DWLW--PAIMML----- 166
 DB 192 MHAHAFKIHGVPRFTPLQPLTIDRRSGLLIPSAIGSSQGLMTAODIYNNIAPNYDL 251
 QY 167 --PDWYGGWPRSGEIDIETIGNRDFKNTGSEFLGIQKMGSTHMGPGMDNRWLT 223
 DB 252 TETPRYMSRQWQANGEFRLYSIGE--GKVAGEYLG-----KRYSEVSDNR----- 298
 QY 224 SLPKSDDMNNGDNHTHTFWMSPNGLRPFVDENQALLDVPRPLIDANPMWDFWE-WG 282
 DB 299 --KRHLFTYNNHNSFLQNW-----RINIVTRYSDKRYNDFDSIYG 338
 QY 283 KP--WLPOYENDPMAGTNLAPDQNFHILINAVAGGTNGF-IPDGCINNGDPA--- 335
 DB 339 RSTGDIANDY-----ARAIYQPNRNFELS-----AHQFQIFDDIVNIGPRYAVPQ 384
 QY 336 -----LQKPMNSGDMWYNDAMRKFPD 355
 DB 385 LDFNYHKYDLANGWMLNFKLHSQAVR--FD 411

RESULT 13
 XYND_RUMFL STANDARD; PRT; 802 AA.
 AC Q53117;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE xylanase/beta-glucanase precursor [includes: Endo-1,4-beta-xylanase
 DE (EC 3.2.1.8) (xylanase); Endo-beta-1,3-1,4 glucanase (EC 3.2.1.73)
 DE (1,3-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase)].
 GN XYND.
 OS Ruminococcus flavefaciens.
 CC Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae;
 CC Ruminococcus.
 OC NCBI_TaxID=1265;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=17;
 RX MEDLINE=93259938; PubMed=8491715;
 RA Flint H.J., Martin J., McPherson C.A., Daniel A.S., Zhang J.-X.,
 RA Flint H.J., Martin J., McPherson C.A., Daniel A.S., Zhang J.-X.,
 RA "A bifunctional enzyme, with separate xylanase and beta(1,3-1,4)-

```
Db      763 TVDDMLKAFNGRTEPLFAHYQWVYTNKNKGYSHSGGQNPW 801
|       | :   |   |   |   |   |   |   |   |   |   |
RESULT_14
ID     GDB_BACAM          STANDARD;             PRT;    239 AA.
AC      P07980;
DT      01-AUG-1988 (Rel. 08, Created)
DT      01-AUG-1988 (Rel. 08, Last sequence update)
DT      01-JUN-1994 (Rel. 29, Last annotation update)
DE      beta-glucanase precursor (EC 3.2.1.73) (Endo-beta-1,3-1,4 glucanase)
DE      (1,3'-1,4-beta-D-glucan 4-glucanohydrolase) (Lichenase).
GN      BGJA.
OS      Bacillus amyloliquefaciens.
OC      Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
CX      NCBI_TaxID=1390;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-BE 20/78;
RX      MEDLINE=87192007; PubMed=1106158;
RA      Homeister J., Kurtz A., Borris R., Knowles J.;
RT      "The beta-glucanase gene from Bacillus amyloliquefaciens shows
RL      extensive homology with that of Bacillus subtilis.";
RL      Gene 49:177-187(1986).
CC      -1 CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC      -1 In beta-D-glucans containing 1,3- and 1,4-bonds.
CC      -1 MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC      SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC      -1 SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on ways
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isdb-sib.ch/announce/
CC      or send an email to license@lsb-sib.ch). CC
CC      -----
DR      EMBL; M15674; AAA87323.1; .
DR      PIR; A29091; A29091.
DR      HSSP; P27051; IGBG.
DR      InterPro: IPR000757; Glyco_hydro.16.
DR      Pfam; PF00722; Glyco_hydro_16; 1.
DR      PRINTS; PR00737; GHYDRASE16.
DR      PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
KW      Hydrolase; Glycosidase; Signal.
FT      SIGNAL              1           25
FT      CHAIN               26         239        BETA-GLUCANASE.
FT      ACT_SITE            134         134        NUCLEOPHILE (BY SIMILARITY).
FT      DISULFID            57           86        BY SIMILARITY.
SQ      SSOURCE             239 AA; 26928 MW; A76A64268A7AAAOB CRC64;

Query Match                      4.7%; Score 104.5; DB 1; Length 239;
Best Local Similarity 20.8%; Pred. No. 0.13;
Matches 53; Conservative 31; Mismatches 100; Indels 71; Gaps 12;

OY      71 LEIKPELLADINIPQTGAPEFGTDI-KYNGLVDYVAMTAGCTNTDNNGCYRTGAAGDIIPA 129
|       ||:  ::::|||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      12 LFMSLCIGITSSVAQRGGSEFPFNSTNSGL-WQKADGNSCGDMFC--TWRAINVSMT 67
OY      130 MSNRVR-----TFQKYSLFHGVVVYAHAKRPVCDWLMPALIMLPEDMVYGGWRPSRGIDI 183
|       |::::::::::|FOKYSLFHGVVVYAHAKRPVCDWLMPALIMLPEDMVYGGWRPSRGIDI 183
DB      68 SLCEMRALATSPSYNFFDCGENRSV-----QTYGYG----- 98
OY      184 IEFIGNRDFKNTGGEFLGIQKMSSTMHWGMGD-----NRMYLTSLPKHS 229
|       |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      99 LYEVRRKKPAKNTGITVSEFFLYTGPTI--GRPMWEIDIEFIKGTTKVQFNYIYINGAGNH 156
OY      230 DDMNYC----DNHTTWFDMSPPGLAREFYDDE--NQALLDVYPPLIDANPWWDVFWENG- 282
|       ::::|::|  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      157 KPFADLGFDDANANAHTYAADFWMQPNISIKWYVDGQLKHATATQIP-----AAPGKTMMNLWG 211
```

QY 283 ---KRWLPQYENDNP 294
 Db 212 TGVDDMLGSYNGVNP 226

RESULT 15

KRE6_CANAL STANDARD: PRT: 740 AA.

ID KRE6_CANAL
 AC P87023;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Beta-glucan synthesis-associated protein KRE6.
 GN KRE6.
 OS Candida albicans (Yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; mtosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=34476;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97234650; PubMed=9079924;
 RA Mio T., Yamada-Okabe T., Yabe T., Nakajima T., Aisawa M.,
 RA Yamada-Okabe H.;

RT "Isolation of the Candida albicans homologs of Saccharomyces
 cerevisiae KRE6 and SKN1: expression and physiological function.";
 RL J. Bacteriol. 179:2363-2372(1997).
 CC -1- FUNCTION: INVOLVED IN THE SYNTHESIS OF (1->6)- AND (1->3)-BETA-D-
 GLUCAN POLYMERS OF THE YEAST CELL WALL IN VIVO. IT IS REQUIRED
 FOR FULL ACTIVITY OF BETA-GLUCAN SYNTHASE IN VITRO. IT MAY BE A
 BETA-GLUCAN SYNTHASE, PART OF A MULTIPROTEIN GLUCAN SYNTHASE OR
 A MODULATOR.

CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Golgi.
 CC -1- SIMILARITY: STRONG, TO SKN1.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 or send an email to license@sdb.ch).

CC
 DR EMBL; D88490; BAA19593.1; -
 DR InterPro; IPR005629; DUF338.

DR Pfam; PF03935; DUF338; 1.
 KM Glycoprotein; Transmembrane; Cell wall; Golgi stack; Signal-anchor.

FT DOMAIN 1 275 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 276 296 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT (POTENTIAL).

FT DOMAIN 297 740 LUMENAL (POTENTIAL).
 FT DOMAIN 40 46 POLY-GLU.
 FT DOMAIN 49 56 POLY-SER.
 FT DOMAIN 126 134 POLY-ASN.
 FT DOMAIN 141 147 POLY-GLY.
 FT DOMAIN 198 201 POLY-SER.

FT CARBOHYD 399 399 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 486 486 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 568 568 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 593 593 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 721 721 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT SEQUENCE 740 AA; 82456 MW; 9DA0537F0B357840 CRC64;

Query Match 4.6%; Score 103.5; DB 1; Length 740;
 Best Local Similarity 17.1%; Pred. No. 0.54;
 Matches 78; Conservative 48; Mismatches 126; Indels 203; Gaps 19;

QY 20 DMDQYHYVODEPFDGAKMOHEVTATGGNSEFOLYTDGANSFVRDGLKLEIKPTLLA 79
 Db 347 DGEWVLVFSDEFD-----AEGRITYESGDQDEFTAP----- 377

QY 80 DNINPQTGAPFGTDMFNGVLDV-WAMTGACNT-----DNNGCYRTGAAGDI 126
 Db 80 DNINPQTGAPFGTDMFNGVLDV-WAMTGACNT-----DNNGCYRTGAAGDI 126

Db 378 -----DIHYDATKDLWYDPDAVTTANGTLNLRMDAYRNHNLFYRSG----- 419
 QY 127 PRAMSAVRYTPQKYSFTSHRNVYHAKMP-VGD--WLMPAIMLPE-----DWYVGMP 176
 Db 420 -----WVOSMQLCTTQGHLEISARLPYGNVTLGMLSMGNLGRPGYLGSTDSGVP 473
 QY 177 RS-----GEIDIE----- 185
 Db 474 YVDSQDAGITPNQSSPDGISYLPQRLNKCCTPGELHPRRGVGRGAPEDVIEGEVMTD 533
 QY 186 TTGNRDFKNTGGEFLGIQKMG-----STMHGPGWDNRKYLTLSPKH 228
 Db 534 SSGKKEKCVASQSLAPMDIWTYIPDYMWVEIYNFSYSTMTNTYTGPPQOALSATTMLN 593
 QY 229 SDDWNYGDNFHTFWPWSNGLRFYVDENQALDVPYPLIDANPMWVDEWEGKPLPQ 288
 Db 594 VTWYEFGDNAHNP-----QTYGYEYLDNPETGYL-----RNFV----- 626
 QY 289 YENDNFWAGTNLAPDQNFHFLNVAVGVTNGFIPDGCINRGDPAOKPWSNGDWYND 348
 Db 627 --GDDP-----TLTYVSQLH-----PDG--NIGWRPLSKREPM----- 656
 QY 349 AMRKEFDARGNMKW-TWDEGDNNAMQVDIIRYK 382
 Db 657 -LILNLGISNNMAYIDMPSISFVTFRIDYVRYQ 690

Search completed: September 16, 2003, 11:26:31
 Job time : 30.0831 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2003, 11:22:17 ; Search time 118.005 Seconds

(without alignments)
839.729 Million cell updates/sec

Title: US-09-596-101c-3

Perfect score: 2240
Sequence: 1 MRRWLVVLCLEFGEFAFTD.....DDEGNNAMQVDYIRYKRN 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL.23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2240	100.0	384	5	077072	077072 Eisenia foe
2	2023	90.3	382	5	095VY3	095VY3 lumbricus t
3	787	35.1	376	5	08MVS9	08MVS9 litopenaeus t
4	781	34.9	366	5	08N0N3	08N0N3 penaeus mon
5	716.5	32.0	361	5	0900G4	0900G4 pacifastacu
6	713	31.8	499	5	026660	026660 strongyloce
7	703	31.4	395	5	017492	017492 anopheles g
8	601	26.8	462	3	09C236	09C236 neurospora
9	596.5	26.6	371	5	08T9V2	08T9V2 aedes aegypt
10	417	18.6	533	16	08CWM7	08CWM7 vibrio vuln
11	413.5	18.5	278	3	096T05	096T05 alternaria
12	390.5	17.4	495	5	09N189	09N189 bombyx mori
13	369.5	16.5	877	2	045095	045095 bacillus cl
14	357	15.9	492	5	09VVR5	09VVR5 drosophila
15	353	15.8	1321	2	059328	059328 clostridium
16	350	15.6	276	2	052734	052734 rhodotherm

17	350	15.6	301	16	09AB54	09AB54 caulobacter
18	350	15.6	435	2	06B641	06B641 oerskovia x
19	340	15.2	482	5	08ISB6	08ISB6 manduca sex
20	333.5	14.9	565	2	08GRB4	08GRB4 pseudomonas
21	329.5	14.7	487	5	09N198	09N198 manduca sex
22	326.5	14.6	488	5	08M0J5	08M0J5 plodia inte
23	323	14.4	494	5	09NHB0	09NHB0 drosophila
24	314	14.0	467	5	017233	017233 bombyx mori
25	312.5	14.0	411	2	08KH33	08KH33 bacillus cl
26	307	13.7	306	2	051333	051333 oerskovia x
27	302	13.5	481	5	096363	096363 hyphantria
28	295	13.2	490	5	09VSR4	09VSR4 drosophila
29	295	13.2	490	5	09NHA8	09NHA8 drosophila
30	290.5	13.0	752	2	08GRB5	08GRB5 pseudomonas
31	287	12.8	646	2	060039	060039 thermotoga
32	286.5	12.8	422	2	09ZG90	09ZG90 flavobacter
33	284	12.7	642	16	09RXN1	09RXN1 thermotoga
34	275.5	12.3	285	16	09F3A0	09F3A0 streptomyces
35	263	11.7	458	16	08PM20	08PM20 xanthomonas
36	262.5	11.7	297	17	08U4K8	08U4K8 pyrococcus
37	257.5	11.5	297	17	073951	073951 pyrococcus
38	249.5	11.1	410	5	08IOR6	08IOR6 drosophila
39	249.5	11.1	432	5	08IOR7	08IOR7 drosophila
40	249.5	11.1	461	5	09VVR4	09VVR4 drosophila
41	248.5	11.1	452	16	08PBD4	08PBD4 xanthomonas
42	247	11.0	383	2	08GCZ5	08GCZ5 lysobacter
43	240	10.7	254	2	08GCZ7	08GCZ7 lysobacter
44	239	10.7	422	16	09EWR5	09EWR5 streptomyces
45	209	9.3	673	5	027082	027082 tachypleus

ALIGNMENTS

RESULT 1

ID 077072 PRELIMINARY; PRT: 384 AA.
AC 077072;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE Coelomic cytolytic factor 1.
GN Ccfl.
OS Eisenia foetida (Common brandling worm). (Common dung-worm).
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC Lumbricina; Lumbricidae; Eisenia.
CX NCBI_TaxID-6396;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98406152; PubMed=9733802;
RA Besschn A., Bilej M., Hanssens F., Raymakers J., Van Dyck E.,
RA Reverts H., Brys L., Gomez J., De Baetselier P., Timmermans M.,
RT "Identification and cloning of a glucan- and lipopolysaccharide-
RT binding protein from Eisenia foetida earthworm involved in the
RT activation of prophenoloxidase cascade.";
RL J. Biol. Chem. 273:24948-24954(1998).
DR EMBL: AF030028; AAC35887.1; -;
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00722; Glyco_hydro_16; 1.
SQ SEQUENCE 384 AA; 44322 MW; C90B5C94003BAD6D CRC64;

Query Match 100.0%; Score 2240; DB 5; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.9e-167;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRRWLVVLCLEFGEFAFTDQYHIYWDQEFYFDGAKWQHEVTATGSGNSEFQLYTQD 60
DB 1 MRRWLVVLCLEFGEFAFTDQYHIYWDQEFYFDGAKWQHEVTATGSGNSEFQLYTQD 60
QY 61 GANSFVBDGKLFKPLLDNINPQNGARPGTDFMNGVLDVAMAYGACNTDNNGCYRT 120
DB 61 GANSFVBDGKLFKPLLDNINPQNGARPGTDFMNGVLDVAMAYGACNTDNNGCYRT 120

```
QY 121 GAAGDIPAMSAARVTFQKYSFTHGRVYVHAKMPVGDMLPMLIMLIPEDVYVYGGWPRSGE 180
DB 121 GAAGDIPAMSAARVTFQKYSFTHGRVYVHAKMPVGDMLPMLIMLIPEDVYVYGGWPRSGE 180
QY 181 IDIETIGNRDPKNTGGFELGIQKMGSTHMGPGMDNRWYLTSLPKHSDDMNYGDNFHT 240
DB 181 IDIETIGNRDPKNTGGFELGIQKMGSTHMGPGMDNRWYLTSLPKHSDDMNYGDNFHT 240
QY 241 FWFDMSPNGLRFFVDDENQALLDVPYPLIDANPMWVDFEMWCKPMLPOYENDNPMWAGTN 300
DB 241 FWFDMSPNGLRFFVDDENQALLDVPYPLIDANPMWVDFEMWCKPMLPOYENDNPMWAGTN 300
QY 301 LAPFDQNFHFLINAVAGGTNGEIPDGCINRGDPAALQKPSNGDWYNDAMRKFFDARGNM 360
DB 301 LAPFDQNFHFLINAVAGGTNGEIPDGCINRGDPAALQKPSNGDWYNDAMRKFFDARGNM 360
QY 361 KWTWDEGDNNAQVDIRVYKRN 384
DB 361 KWTWDEGDNNAQVDIRVYKRN 384

RESULT 2
Q95VY3 PRELIMINARY; PRT; 382 AA.
ID 095VY3:
AC 095VY3:
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 21, last annotation update)
DE Coelomic cytolytic factor precursor.
OS Lumbriculus terrestris (Common earthworm).
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haploclaxida;
OC Lumbricina; Lumbricidae; Lumbriculus.
OX NCBI_TaxID=6398;
RN [1]
RP SEQUENCE FROM N.A.
RA Beschin A., De Baetselier P., Bilek M.;
RT "Distinct carbohydrate recognition domains of an earthworm defense
RT molecule recognize Gram negative and Gram positive bacteria.";
RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF395805; AAL09587.1; -.
DR InterPro; IPR00757; Glyco_hydro.16.
DR Pfam; PF00722; Glyco_hydro.16; 1.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 382 AA; 43931 MW; 5256CF171EB7D3FB CRC64;

Query Match 90.3%; Score 2023; DB 5; Length 382;
Best Local Similarity 90.1%; Pred. No. 1.9e-150;
Matches 346; Conservative 15; Mismatches 21; Indels 2; Gaps 1;

QY 1 MRWTLVYICLFGEGFAITDMQVHIWODEFDYFDGAKWOHEVYATGGGSEFQLYTOD 60
DB 1 MRWTLVYICLFGEGFAITDMQVHIWODEFDYFDGAKWOHEVYATGGGSEFQLYTOD 60
QY 61 GANSEVRGKLEFKPTLLADININPOTGAFEGTDFMYNGVLDVWAMYAGACTINDNNGCYRT 120
DB 61 GANSEVRGKLEFKPTLLADININPOTGAFEGTDFMYNGVLDVWAMYAGACTINDNNGCYRT 120
QY 61 SRNSVVRGKLEFKPTLLADININPOTGAFEGTDFMYNGVLDVWAMYAGACTINDNNGCYRT 120
DB 61 SRNSVVRGKLEFKPTLLADININPOTGAFEGTDFMYNGVLDVWAMYAGACTINDNNGCYRT 120
QY 121 GAAGDIPAMSAARVTFQKYSFTHGRVYVHAKMPVGDMLPMLIMLIPEDVYVYGGWPRSGE 180
DB 121 GAAGDIPAMSAARVTFQKYSFTHGRVYVHAKMPVGDMLPMLIMLIPEDVYVYGGWPRSGE 180
QY 181 IDIETIGNRDPKNTGGFELGIQKMGSTHMGPGMDNRWYLTSLPKHSDDMNYGDNFHT 240
DB 181 IDIETIGNRDPKNTGGFELGIQKMGSTHMGPGMDNRWYLTSLPKHSDDMNYGDNFHT 240
QY 241 FWFDMSPNGLRFFVDDENQALLDVPYPLIDANPMWVDFEMWCKPMLPOYENDNPMWAGTN 300
DB 241 FWFDMSPNGLRFFVDDENQALLDVPYPLIDANPMWVDFEMWCKPMLPOYENDNPMWAGTN 300
QY 301 LAPFDQNFHFLINAVAGGTNGEIPDGCINRGDPAALQKPSNGDWYNDAMRKFFDARGNM 360
DB 301 LAPFDQNFHFLINAVAGGTNGEIPDGCINRGDPAALQKPSNGDWYNDAMRKFFDARGNM 360
QY 361 KWTWDEGDNNAQVDIRVYKRN 384
DB 361 KWTWDEGDNNAQVDIRVYKRN 384
```

```
QY 361 KWTWDEGDNNAQVDIRVYKRN 384
DB 361 KWTWDEGDNNAQVDIRVYKRN 384

RESULT 3
Q8MVS9 PRELIMINARY; PRT; 376 AA.
ID 08MVS9:
AC 08MVS9:
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, last annotation update)
DE Lipopolysaccharide and beta-1,3-glucan binding protein.
OS Lipopneustes stylirostris (blue shrimp).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
OC Penaeidae; Litopenaeus.
OX NCBI_TaxID=29019;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=hepatopancreas;
RX MEDLINE=2068042; PubMed=12072514;
RA Roux M.M., Pain A., Klimpel K.R., Dhar A.K.;
RT "The lipopolysaccharide and beta-1,3-glucan Binding Protein Gene Is
RT stylirostris";
RL J. Virol. 76:7140-7149(2002).
DR EMBL; AF473579; AAM73871.1; -.
DR InterPro; IPR00757; Glyco_hydro.16.
DR Pfam; PF00722; Glyco_hydro.16; 1.
SQ SEQUENCE 376 AA; 42610 MW; D7ADC5A10208885 CRC64;

Query Match 35.1%; Score 787; DB 5; Length 376;
Best Local Similarity 43.0%; Pred. No. 1.2e-53;
Matches 157; Conservative 58; Mismatches 106; Indels 44; Gaps 11;

QY 26 IWODEFEDGAKWOHEVYATGGGSEFQLYTODGANSFVRGKLEFKPTLLADININPQ 85
DB 26 IWODEFEDGAKWOHEVYATGGGSEFQLYTODGANSFVRGKLEFKPTLLADININPQ 85
QY 42 MIFEDNFYLDNDVMEHELTMSGGWEEQAYNNNSISTRTSTLFIKRELTA----- 96
DB 42 MIFEDNFYLDNDVMEHELTMSGGWEEQAYNNNSISTRTSTLFIKRELTA----- 96
QY 86 TGAPEGTDFMYNGVLDVWAMYG---ACTNTDNNNGCYRTGAAGD-IPAMSAARVTFQKYS 141
DB 86 TGAPEGTDFMYNGVLDVWAMYG---ACTNTDNNNGCYRTGAAGD-IPAMSAARVTFQKYS 141
QY 97 ---WAGDDFLTSGTIDLMGNGRGDVCYCTGSSYSSGSSNLVNPVLSARLTMSNFA 153
DB 97 ---WAGDDFLTSGTIDLMGNGRGDVCYCTGSSYSSGSSNLVNPVLSARLTMSNFA 153
QY 142 FTNHRVYVHAKMPVGDMLPMLIMLIPEDVYVYGGWPRSGEIDIETIGNRDPKNTGGFELG 201
DB 142 FTNHRVYVHAKMPVGDMLPMLIMLIPEDVYVYGGWPRSGEIDIETIGNRDPKNTGGFELG 201
QY 154 FRYGRLEIRAKKPRGDMPLPMLPRNMPYGAMPASGEIDLESKNDNFGT-----LG 208
DB 154 FRYGRLEIRAKKPRGDMPLPMLPRNMPYGAMPASGEIDLESKNDNFGT-----LG 208
QY 202 IOKMGSTHMGPGMDNRWYLTSLPKHSDDMNYGDNFHTFWFDMSPNGLRFFVDDENQAL 261
DB 202 IOKMGSTHMGPGMDNRWYLTSLPKHSDDMNYGDNFHTFWFDMSPNGLRFFVDDENQAL 261
QY 262 LDVPPYPLIDANPMWVDFEMWCKPMLPOYENDNPMWAGTNLAPFDQNFHFLINAVAGTNG 321
DB 262 LDVPPYPLIDANPMWVDFEMWCKPMLPOYENDNPMWAGTNLAPFDQNFHFLINAVAGTNG 321
QY 269 ID-----PGNSFW-DESGMDSVY-----DNPMSAGSKMAPDQKFFYLLINAVAGTNG 315
DB 269 ID-----PGNSFW-DESGMDSVY-----DNPMSAGSKMAPDQKFFYLLINAVAGTNG 315
QY 322 FIPDGCINRGDPAALQKPSNGDWYNDAMRKFFDARGNMKNTWD-DEG---DNNNAQVYD 377
DB 322 FIPDGCINRGDPAALQKPSNGDWYNDAMRKFFDARGNMKNTWD-DEG---DNNNAQVYD 377
QY 316 FFP-----DDVSKPWSN--LSPTAFLDFWNARDEMLPSWQAGEGRISGAAMQVYD 365
DB 316 FFP-----DDVSKPWSN--LSPTAFLDFWNARDEMLPSWQAGEGRISGAAMQVYD 365
QY 378 IRYK 382
DB 366 VRYK 370

RESULT 4
Q8N0N3 PRELIMINARY; PRT; 366 AA.
ID 08N0N3:
AC 08N0N3:
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, last annotation update)
```

```

DE Beta-1,3-glucan binding protein.
OS Peneus monodon (Penaeid shrimp).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
OC Penaeidae; Peneus.
OX NCBI_TaxID=6687;
RN [1]
RP SEQUENCE FROM N.A.
RA Sittunvalucksana K., Lee S.Y., Soderhall K.;
RT "The beta-1,3-glucan binding protein from the black tiger shrimp,
RT Peneus monodon."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF366168; AAM21213.1; -
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00722; Glyco_hydro_16; 1.
SQ SEQUENCE 366 AA; 41497 MW; 6F540A0E83FDD7D CRC64;

Query Match          34.9%; Score 781; DB 5; Length 366;
Best Local Similarity 43.3%; Pred. No. 3.3e-53;
Matches 159; Conservative 50; Mismatches 110; Indels 48; Gaps 10;

QY 26 IWODEFDYFDGAKMOHEVTATGGNSEPOLYTODGANSFVRDGLFIKPTLLADNINPQ 85
DB 32 MIFEDNFQYLDNDIMEHEITMSGGWMEEQAYVNNNSISYTRSTLFIKPDL----- 83
QY 86 TGAPFGTDPMYNGVLDVWAMYG---ACTTDDNNGCYRTGAAGD-IPPAMSAVRYTOKYS 141
DB 84 TSNMKGEDFLSGTGLDMGNRGDVCYTGNSYYGSGSRVSSSNIIINPVSARLITMSNFA 143
QY 142 FTHGRVYVAKKPVGDMLPAILMPLPDMVYGGMPRSGSEIDITETIGNRDFKMTGSEFLG 201
DB 144 FRYGLLEVAKKPRGDMPLPAILMPLPDMVYGGMPRSGSEIDITETIGNRDFKMTGSEFLG 201
QY 202 IOKMGSMTMHGPGMDNRYWLSLPKHSDDWNYGDNFHTFWMDSNGRLFFVDDENQAL 261
DB 199 NQYGGFTLHMCGFMPNFEKTHAEYSANTGSPADDFHWRLDWTQDNMEFYVDVLT 258
QY 262 LDVPPPLIDANWVWDFWEMG--KPYLPQYENDNPAGGTNLAPFQDNHFLINAVGCT 319
DB 259 VDP-----GSEWDFAGNGPFP-----DNPWAAAGAKAPFQDKFYLLINAVGGT 303
QY 330 NGFIPDGCINRGDPALQKPMNSGDWYNDAMRKFFDARGNMWKTW----DDEGDNNAMOV 375
DB 304 NGFFPDGLAS-----KPMNSN--LSPTAFIDFWARNRDBMLPSMAGEDRISGAAAOV 353
QY 376 DTRVYK 382
DB 354 DYVRVWK 360

RESULT 5
QY00G4 PRELIMINARY; PRT; 361 AA.
AC QY00G4;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
DE Lipopolysaccharide and beta-1,3-glucan binding protein precursor.
GN LGBP.
OS Pacifastacus leniusculus (Signal crayfish).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidae;
OC Astacoidea; Astacidae; Pacifastacus.
OX NCBI_TaxID=6720;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-Hemocyte;
RC MEDLINE-20092910; PubMed-10625682;
RX Lee S., Wang K., Soderhall K.;
RA "A lipopolysaccharide- and beta-1,3-glucan-binding protein from
RT hemocytes of the freshwater crayfish Pacifastacus leniusculus:
RT purification, characterization, and cDNA cloning."
RL J. Biol. Chem. 275:1337-1343(2000).

```

```

DR EMBL: AJ250128; CAB65353.1; -
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00722; Glyco_hydro_16; 1.
KW Signal.
FT SIGNAL
FT CHAIN
FT 1 15
FT 16 361
FT 361 15
SQ SEQUENCE 361 AA; 41043 MW; 93A5E67911ED6619 CRC64;

Query Match          32.0%; Score 716.5; DB 5; Length 361;
Best Local Similarity 40.9%; Pred. No. 3.6e-48;
Matches 148; Conservative 68; Mismatches 105; Indels 41; Gaps 12;

QY 26 IWODEFDYFDGAKMOHEVTATGGNSEPOLYTODGANSFVRDGLFIKPTLLADNINPQ 85
DB 30 LIFNDQFNDLNRWVMPREVTMSGGWMEEQAYVNNNSISYTRSTLFIKREL----- 81
QY 86 TGAPFGTDPMYNGVLDVWAMYGACTTDDNNGCYRTGAAGD-IPPAMSAVRYTOKYSFTH 144
DB 82 TSKWYSEHFLFNDELN--LGDKCTDHRDYGCVKRGSEHIIINPMSAKFTTHPSFAFRY 138
QY 145 GRVYVAKKPVGDMLPAILMPLPDMVYGGMPRSGSEIDITETIGNRDFKMTGSEFLG 204
DB 139 GRVYVAKKPVGDMLPAILMPLPDMVYGGMPRSGSEIDITETIGNRDFKMTGSEFLG 204
QY 205 MGSMTMHGPGMDNRYWLSLPKHSDDWNYGDNFHTFWMDSNGRLFFVDDENQALDV 264
DB 194 AGSTLHMGPVPAQNMFLKTHKTYSSANDGSPANNFHIWRMDTRDNMKEFYDDQLVDP 253
QY 265 PYYLIDANWVWDFWEMGKRWLPQYENDNPAGGTNLAPFQDNHFLINAVGCTNCFIP 324
DB 254 -----GSEWDFAGNGPFP-----GLGNSON-NPWRGSKAPPDQKFFYLLINAVGCTNGYFP 300
QY 325 DGCINRGDPALQKPMNSGDWYNDAMRKFFDARGNMWKTW--DEG--DNNAMOVDIRV 380
DB 301 DGV---SNPA--KPMNNNSPH--ASRDPMNAGSWLPSEHGEHISENALKVDIVKV 353
QY 381 YK 382
DB 354 WK 355

RESULT 6
QY0660 PRELIMINARY; PRT; 499 AA.
AC QY0660;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Beta 1,3-glucanase.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinozoa; Echinodermata; Echinozoa; Echinozoa; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=668;
RN [1]
RP SEQUENCE FROM N.A.
RA Bachman E.S., McClay D.R.;
RT "Molecular cloning of the first metazoan beta-1,3 glucanase from eggs
RT of the sea urchin Strongylocentrotus purpuratus."
RL Proc. Natl. Acad. Sci. U.S.A. 93:6808-6813(1996).
DR EMBL: U49711; AAC47235.1; -
SQ SEQUENCE 499 AA; 55275 MW; D863F336BEOCF5AC CRC64;

Query Match          31.8%; Score 713; DB 5; Length 499;
Best Local Similarity 42.7%; Pred. No. 1e-47;
Matches 158; Conservative 48; Mismatches 88; Indels 76; Gaps 15;

QY 26 IWODEFDYFDGAKMOHEVTATGGNSEPOLYTODGANSFVRDGLFIKPTLLADNINPQ 85
DB 179 LIFQEFDSFNLDIMEHEITAGGGWMEEQAYVNNNSISYTRSTLFIKPTLLTDDKL--- 235

```



```

Db 366 AEMRENDTLLPENPAGSDSTGNAPEFDPFYLLINVAAGSRIGWFPD---NKG-----DK 417
QY 339 PW----SNGDWYNDAARKPFEDARGMKMKWDEGDNNAQVYIRYKR 383
Db 418 PMLDSATNAQW-----TFMSAADQMLPTW-GEQDQGMATVSKVMRQ 459

RESULT 9
ID 08T9V2 PRELIMINARY; PRT: 371 AA.
AC 08T9V2.
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Putative secreted protein.
OS Aedes aegypti (Yellowfever mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.
OX NCBI_TaxID=7159;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Black eye; TISSUE=Salivary gland;
RA Valenzuela J.G., Pham V.M., Garfield M.K., Francischetti I.M.,
RA Ribeiro J.M.C.;
RT "Toward the stailome of the adult female mosquito Aedes aegypti.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF46594; AAL76017.1;
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00722; Glyco_hydro_16; 1.
SQ SEQUENCE 371 AA; 41943 MW; 1FE2E5DC7B665CE6 CRC64;

Query Match 26.6%; Score 596.5; DB 5; Length 371;
Best Local Similarity 35.4%; Pred. No. 9.4e-39;
Matches 129; Conservative 71; Mismatches 115; Indels 49; Gaps 13;

QY 26 IWMODEDIDYDGAQWQHEVTATGSGNSEFQLYTDGANSFVRDGLFKPTLLADNINPQ 85
Db 47 LITQDNFNRLDRNVMQHEHNSIGGCGNNEFQWYSGSGRNSYIKNNMLYIRPTLTSD----- 102
QY 86 TGAPEGDFMYNGVLDV--NAMYGACTNTDN-----NCTYRTGAAGDI--PPAMARVTF 137
Db 103 -----TGAFALSGYGININEGPOSQCTDAPGMAQOIHCCYRRGSDRIILNIVRSARLRTV 158
QY 138 QKTSFTGRVYVHAHAKPFVGDMLPAIMLPEPDWYGGPFRSGEIDITETIGNRDKNNG 197
Db 159 NSFAPFKGKYEINKLQGDMLPALMLLPKGDYTGTPKSGEVDLMSRGNRLNVO--NN 217
QY 198 EPLGIQMGSTMHMGPMWDNRKYLTLPHKSHDDMNTGDNFTWDFWSPNGLRFFVDDE 257
Db 218 EKIGIQKVSCLHFG-----DNPNVRSSQCGSVSGNLFGAMFNRRQLWTWKVIGQINDR 273
QY 258 NQALLDVPYPLIDANPFWVDFEWGKFWLPDYENDNPWAGSTNLAPEDQNFHILNVAVG 317
Db 274 IFTTV--TPYE-----GFWRLLGG-----FSFNFWPKSKKAPEDKEFYIWMNVAVG 317
QY 318 GTNGFLPDGGINRGDPALQKPSNGDMYNDAMRKFPDARKNMKTWDDSDNNAMQVYD 377
Db 318 G--DYFPDAMN-----PHKPKMRQGN--PSAMTDFYKASNMVSTW--GDAALALEVDW 365
QY 378 IRVY 381
Db 366 VKVW 369

RESULT 10
ID 08CMT7 PRELIMINARY; PRT: 533 AA.
AC 08CMT7.
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Beta-glucanase/beta-glucan synthetase.
GN VV21300.

```

```

OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhoe J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF016812; AAO08191.1;
KW Complete proteome.
SQ SEQUENCE 533 AA; 58976 MW; E9B2B914FA744DD1 CRC64;

Query Match 18.6%; Score 417; DB 16; Length 533;
Best Local Similarity 29.1%; Pred. No. 1.7e-24;
Matches 118; Conservative 45; Mismatches 125; Indels 118; Gaps 13;

QY 5 LVVCLLFGEGFAFTDW-----DOYHIWQDEF--DYFDGAKWQHEVT 45
Db 20 LTVGQSOTGDSASATDLYQTKKRVMLQDQPTPSPDQMLVWVDFDQDKINKRMVLEEN 79
QY 46 ATGGGSEFQLYTDGANSFVRDGLFT---KPTLLA-DNINQTAGPCTDPMYGVLD 101
Db 80 CMGGGNEQCCYTKRRARNAFVODGYLIVAHHEESYSGPNPECKVGA----- 126
QY 102 VWAMYGACTNTDNGCYRTGAAGDI--PPAMARVRFQKYSFTGRVYVHAHAKPFVGDMLP 161
Db 127 -----GANKTLPFSARLRTGKKDKHDKRGREIRAKLPSGGCTWP 166
QY 162 AIWMLPEPDWYGGWPRSGEIDITETIGNDFKNTGGEFT--GIQKMGSTMHMGPMWDNR 219
Db 167 AIWMLPPTNKYGTWMAASGEIDIMEAVNLKAQSDAPGAQAGDGENRITYGSLHYGKAMPDV 226
QY 220 Y--WLTSLPHKSHDDMNTGDNFTWDFWSPNGLRFFVDDENQALLDVPYPLIDANPWWVD 277
Db 227 YSGGASLSP--NNINPADDFHTYALWEWEGELRMYVDINHAT----- 267
QY 278 FWEWGRPMLPYENDNPWAGSTNLAPEDQNFHILNVAAGTNGEFLPDGGINRGDPALQ 337
Db 268 --QTDQEWYSQKVKVGDALVNAKGAPEPDRLLHLLNLAAVGS----- 307
QY 338 KPSNGDMYNDAMRKFPDARGMKMKTWDEGD--NNAMQVYIRYVK 362
Db 308 --WS-----ANANQKGI--DKSDFPKTMVLDYKVKYR 335

RESULT 11
ID 096T05 PRELIMINARY; PRT: 278 AA.
AC 096T05.
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Mixed-linked glucanase (Fragment).
OS Alternaria alternata.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
OC Pleosporales; Pleosporaceae; mitosporic Pleosporaceae; Alternaria.
OX NCBI_TaxID=5599;
RN [1]
RP SEQUENCE FROM N.A.
RA Eshel D., Prusky D., Dinnor A.;
RT "Mixed-linked glucanase precursor of Alternaria alternata.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF282319; AAK69516.1;
DR InterPro: IPR000757; Glyco_hydro_16.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
FT NON_TER 1
SQ SEQUENCE 278 AA; 31073 MW; B53AB7749826B22D CRC64;

Query Match 18.5%; Score 413.5; DB 3; Length 278;

```


QY 307 NHEFLINAVAGCT--NGFLPDGGINRGDPAALOKPMSNGDWYNDAMKRFDFDARGNMKWTW 364
 DB 646 PFFILMNLAIGCTFGDGRTPD-----PSDIPA----- 672
 QY 365 DDEGDNNAMOVYIRYK 382
 DB 673 -----TWQGVYRVYK 683

RESULT 14
 09YVR5 PRELIMINARY; PRT; 492 AA.
 AC 09YVR5;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, last annotation update)
 DE CG6895 protein.
 GN CGBP1 OR CG6895.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RX MEDLINE=20196006; PubMed=10731132;
 RC STRAIN-BERKELEY;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abriil J.F., Agbayani A., An H.-J., Andrews-Planckoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spletter E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner A., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang Q., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT *The genome sequence of Drosophila melanogaster.*;
 RT Science 287:2185-2195(2000).
 RL EMBL: AE005515; AAF49244.1; -;
 DR Flybase: FBgn0040323; GDBP1.
 DR InterPro: IPR000757; Glyco_hydro.16.
 DR Pfam: PF00722; Glyco_hydro.16; 1.
 SO SEQUENCE 492 AA; 55314 MW; C5D0E561FAB779 CRC64;

Query Match

15.9%; Score 357; DB 5; Length 492;

Best Local Similarity 27.2%; Pred. No. 7.6e-20;
 Matches 101; Conservative 67; Mismatches 133; Indels 70; Gaps 16;

QY 26 IWVDEEDYDYGAKWQHEV-TATGNGSEPOLYPODANSVRCKLFIKPTLLADNINP 84
 DB 175 LFEETFDQLNESLWIDVRLPLDSKDAEFLV--DG-KAVYHGNLVEP-LTWSSYRP 230
 QY 85 QTGAFEGTDEMYNGLDVMAMYGACTYNDN--NGCY-----RTGAAGDIPAMSAVRFTQ 138
 DB 231 DLS-----IANRRLD---LSERCTGHNKIKECILHSTGSGSGINPPIVTPRISKE 280
 QY 139 KYSTFGRVYVYHAKMPYQWMLPAIMLP-BDWYVYGNPMSGIEDIETIGNRDKNTG 197
 DB 281 TFAFOYGRIRIETRAIPKGMIVPPLLEPLETMYGQSGYEQGLVALARNGSVLRMBRG 340
 QY 198 EFLGICQKMGSMHMGPCWDDNRYLTLSPKISDW-----NNGDNHTTWPMSPBGL 250
 DB 341 KLVD---GNSLVGSP-----VLSTDAHQREDLMLSKRKISHFGDDFTYSLWSSNRL 390
 QY 251 RFVVDENQALDVPYPLIDANPMWVDFWENGRPMLPOYENDNPMAGSTNLAPDONFHF 310
 DB 391 LFSVDQVYGMENLNGFTLENDP-----RWKGGPMAFPDKMYI 430
 QY 311 ILVAVGCTNGFLPDGGINRGDPAALOKPMSNGDWYNDAMKRFDFDARGNMKWTWDEGDN 370
 DB 431 SLGVSVGFGFDF-----VDHLRTATYERKPWAN--YHFOAKLQFHQADQWLPPTWKQP--- 480
 QY 371 NAMOVYIRYK 381
 DB 481 -ALKIDIVRVF 490

RESULT 15
 ID 059328 PRELIMINARY; PRT; 1321 AA.
 AC 059328; 09ALJ34;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)
 DE Endo 1,3(4)-beta-glucanase (EC 3.2.1.6).
 GN IICA
 OS Clostridium thermocellum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OC NCBI_TaxID=1515;
 RX [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN-DSM 1237.
 RA Schwarz W.H.; Schimming S., Fuchs K.P., Staudenbauer W.L.;
 RA Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RC SEQUENCE FROM N.A.
 RC STRAIN-DSM 1237.
 RA Schwarz W.H.;
 RA Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RC SEQUENCE FROM N.A.
 RC STRAIN-DSM 1237.
 RA Zverlov V.;
 RA Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RC SEQUENCE FROM N.A.
 RC STRAIN-F7;
 RA Zverlov V.V., Schwarz W.H.;
 RA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL: X8732; CA61884.2; -;
 DR EMBL: AJ307315; CAC27412.2; -;
 DR HSSP: P14090; IU10.
 DR InterPro: IPR003305; CBM_Cenc.
 DR InterPro: IPR000757; Glyco_hydro.16.
 DR InterPro: IPR001119; SLH.
 DR Pfam: PF02018; CBM_4_9; 4.
 DR Pfam: PF00722; Glyco_hydro.16; 1.

DR Pfam: PF00395; SLH; 3.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 1321 AA: 147772 MW: A45213FF99748F0C CRC64:

Query Match 15.8%; Score 353; DB 2; Length 1321;
Best Local Similarity 27.2%; Pred. No. 5.2e-19;
Matches 100; Conservative 51; Mismatches 93; Indels 124; Gaps 16;

```
QY 22 DQYHYWDEPD--YFDGAKMCHVATATGGNSEFOLYTODGANSFVRDGLFIKPTLLA 79
   ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 423 EEMRLWSEDFNGSEINMANWSYDDPTNGRMNGEVOSTON--NAYIKDALVIEAR--- 477
   ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 80 DNINPQTGAPFGTDFWYNGVLDMVAMYGACTNTDNGCYRTGAAGDIPFAMSAVRFPFK 139
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 478 ---KEDITEPSGETYHYT-----SKLIYTKG 501
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 140 YSFTHGRRVYVHAKMPVGDWLMIPAIMLPEBW-VYGGMPRSGEIDIETIGNRDKNTGGE 198
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 502 KSMKYGKFEIRAKMPQGQIMPAIMMPEDEPFYGTWPKCEIDIMELGHEP----- 554
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 199 FIGIQKMGSTMHGPGCDNRNRYLTLSPKHSDDMNTGDNFHTWDSPNGLRPFVDDEN 258
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 555 ---DKIYGTIHGEPHKESQGYT-LP---EGQTFADDFHYSIEWEPGEIRWYIDGK- 605
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 259 QALLDVPYPLIDANPMWVDFWEMGKPMPLQYENDNPMAGS--TNLAPFDQNFHILNVAY 316
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 606 ---LYHY-----ANDM-----TSRD--PYLADDYTYPAFPDQNFLLNLISV 642
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 317 GGTNGFIPDGCINRGDPALQKPSNGDMYNDAMRKEFFDARGNMKWTWDEGDNNMAQVD 376
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 643 G-----GGWPG-----YPDETYF-----PQGMVVD 663
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 377 YIRVYKRN 384
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 664 YVRVYOKD 671
```

Search completed: September 16, 2003, 11:28:43
Job time : 121.005 secs